A New Perspective on Spatial Heterogeneity in African Development*

Nicolas Suarez
Stanford University

Pascaline Dupas
Princeton University

Zhongyi Tang
Amazon

September 30, 2024

(click here for the most current version)

Abstract

Access to basic infrastructure is a critical component of quality of life and an important measure of economic development. However, on-the-ground data about infrastructure access, especially in low-income countries, is often sparse and costly to collect. We leverage satellite imagery and survey data to train a machine learning model that predict access to infrastructure for each 6.72x6.72km area of Africa. The model achieves accuracy levels of 77.1% to 84.7%.

We show the value of this novel dataset with two applications. First, we use a spatial regression discontinuity design to study how much of the heterogeneity in infrastructure access *across countries* comes from differences in institutional quality, finding a positive effect of a modest magnitude, reconciling previous contradicting results in this literature. Second, we study the role of political favoritism in explaining *within country* heterogeneity, finding that areas with political ties to current or former presidents have better access to infrastructure.

*We are grateful to the Data for Development Initiative at the Stanford King Center on Global Development for funding, attendants of Stanford's third year seminar and development seminar for their feedback, and to Jeremy Weinstein for his participation in early discussions of this project.

1 Introduction

Access to basic infrastructure such as piped water, sewerage, and electricity, is a critical component of quality of life. It contributes to improving health, preventing illness and reducing child mortality (Galiani et al., 2005; Alsan and Goldin, 2019; Irwin et al., 2020). It also helps save time spent on home production, freeing up time for employment (Dinkelman, 2011), school work (Ashraf et al., 2021) or leisure (Devoto et al., 2012). Yet, granular, up-to-date data on access to basic infrastructure is lacking in many low-income countries. Country-level estimates are based on representative household surveys, but those tend to be too sparse to conduct meaningful spatial analyses of how infrastructure access is distributed within a country. This dearth of localized data limits what we know about inequality in access and its determinants.

This paper generates novel data on infrastructure access for every single 6.72km $\times$ 6.72km area of Africa, and uses that data to document some key facts about the role of institutions in the levels and distribution of infrastructure access. We combine information from communities in household surveys and satellite imagery to train a machine learning model following Oshri et al. (2018) to predict infrastructure access (access to piped water, sewerage and electricity). We do so with relatively high precision, achieving accuracy levels of 77.1% to 84.7%. Specifically, 68.4% of patches with sewerage according to ground survey data are predicted to have sewerage, and 90.9% of those without sewerage are predicted not to have sewerage. The figures are 78.6% and 75.6% for piped water and 82.2% and 84.3% for electricity, respectively.

This novel data allows us to revisit some old debates in the political science and economics literature. We first look at the role of national institutions on local development levels. Taking advantage of the fact that national borders that divide African countries are largely the result of boundary-drawing practices from the colonial era and arbitrarily dividing pre-colonial ethnic homelands into different countries (McCauley and Posner, 2015), we employ a spatial regression discontinuity design to study the effect of national institutions on infrastructure development. Results suggest that national institutions over the past 20 years have played some role: crossing from a country with lower institutional quality to a country with higher institutional quality significantly increases the likelihood of access to basic infrastructure. Second, we examine the

allocation of resources within a country. We find a strong positive correlation between political ties and infrastructure access across all infrastructure measures considered: the birthplaces of African presidents have better infrastructure access, though higher institutional quality reduces such political favoritism.

To generate reliable predictions of infrastructure access for Africa, we train a Residual Network (ResNet). This type of Convolutional Neural Network (CNN) architecture has been shown to have good performance for binary classification problems (He et al., 2016). Our ResNet takes six bands from daytime Landsat imagery and 1 band from the VIIRS Nighttime data as input and generates a prediction, then makes a binary classification based on the prediction, while minimizing the loss function between predictions and ground truths. Additionally, we apply a calibration technique, *Temperature Scaling*, so that the scores predicted by the model for a given patch of land can be correctly interpreted as the *probability* that the patch of land has access to the infrastructure (Guo et al., 2017).

We train three models, for the following three types of infrastructure: electricity, sewerage and piped water. Because the models use both daytime and night-time imagery, they outperform predictions based on night lights (NL) alone, in particular for areas far from the capital city, where many areas are completely unlit at night, yet they exhibit important heterogeneity in infrastructure access.

Having generated good granular data on access to infrastructure, we use it to revisit the literature on the determinants of local development. To study the impact of national institutions on economic development, we exploit the arbitrary drawing of national borders during the colonial period for African countries. To do this, we intersect the "Tribal Map" of Murdock (1959) and a map of modern national borders in Africa (GADM, 2018) to identify ethnic groups that were partitioned by national borders, and modern countries that contain adjacent ethnic groups. Then, we use a spatial regression discontinuity design to study the effect of national institutions across borders on access to infrastructure. We find that crossing from a country with lower institutional quality to a country with higher institution quality—corresponding to a 0.77 standard deviation gain in institutional quality on average—increases the likelihood of a patch being lit in nighttime lights (NL) by 3.54 percentage points; and the likelihood of having access to electricity, sewerage and piped water

increases by 1.72, 1.12 and 1 percentage points, respectively. Patches in countries with higher institutional quality also have a paved road 4.73 kilometers closer on average.

We also use our dataset to study the extent to which the allocation of infrastructure across geographical units within a country is consistent with ethnic favoritism. We compute average infrastructure access at the administrative level-2 (typically, the district), and collect information on the birthplace and ethnicity of presidents since independence (or 1960 for countries that were never colonized). Our findings show a strong positive correlation between political ties and infrastructure access across all infrastructure measures considered. The correlation is especially strong for geographical units within the median distance to the capital. By interacting the institutional quality index with the political tie measure, we find that having better institutional quality mitigates the uneven allocation of resources. Lastly, as African countries have been receiving foreign aid in the past decades, we investigate whether there are different patterns of political favoritism in high- vs. low-aid countries. Results demonstrate that the correlation between political ties and uneven allocation is not weaker (and if anything, higher) in high-aid countries. Our paper makes two main contributions. First, we present a new method to generate accurate, hyper-localized data from satellite imagery when granular on-the-ground data is lacking. Such localized data is particularly powerful in investigating spatial heterogeneity. Our proposed methodology can be used to generate similar data for other continents, or other time periods, provided satellite imagery is available and there is at least some ground truth data to train the model.[1] We build on the seminal contributions of Chen and Nordhaus (2011) and Henderson et al. (2012), which introduced night time luminosity measured via satellite as a meaningful proxy for economic development, and the subsequent advances showing how night time luminosity can be combined with on-the-ground surveys to train deep learning models to predict income levels (Jean et al., 2016; Oshri et al., 2018; Ratledge et al., 2022; Xie et al., 2016; Yeh et al., 2020).

Our second contribution is to the debate on the role of institutions in economics development. Prior evidence on this has been mixed: while Acemoglu et al. (2001) estimated a very strong role of institutional quality at the country level using an instrumental variable design,

[1] Our method uses public imagery, which further lowers the barrier in applying our method in other settings.

Michalopoulos and Papaioannou (2014), using a spatial regression discontinuity, failed to detect an effect of national institutions on local African development after controlling for ethnicity fixed effects. Our novel dataset helps to reconcile these two sets of results. Replicating the analysis of Michalopoulos and Papaioannou (2014) using our novel dataset, we find that national institutions do play a role, but the magnitude of their impact is much more modest than that estimated in Acemoglu et al. (2001), confirming the conjecture of Pande and Udry (2005) that the identification strategy in Acemoglu et al. (2001) may have generated an upward bias.

The granular dataset also allows us to revisit the literature on the role of ethnicity in distributive allocations (Bates, 1974). Expanding on the studies of Burgess et al. (2015) on the placement of roads and that of Dreher et al. (2019) on where foreign assistance is spent, we confirm that political favoritism exists across many domains and show that, while it is somewhat exacerbated by foreign assistance, it can be mitigated through higher quality national institutions.

The remainder of the paper proceeds as follows. Section 2 introduces the machine learning model and describes how it is trained and validated, before presenting its predictions for infrastructure access for the African continent, and how we combine these predictions with other outcomes and geographical controls to build the final dataset. Our two political economy applications are presented in Sections 3 and 4. Section 5 concludes.

2 Generating Granular Infrastructure Access Data

Our aim is to generate localized measures of infrastructure access. We do so at the "patch" level. By patch we mean a square geographic area of 6.72 $\times$ 6.72 kilometers. We impose a grid onto the African continent, dividing it into 747,165 patches.

The only measures we are aware of that exist for the entire grid of patches are night lights (from satellite imagery) and population density estimates created through satellite imagery and remote sensing. Our aim is to add measures of infrastructure access, specifically, access to the electricity grid, piped water and the sewerage.

To generate patch-level estimates of infrastructure access, we combine both daylight and

nightlights satellite imagery with deep learning techniques, which have been successfully used in a wide array of satellite imagery applications.[2]

This section first describes the inputs used to train the model (satellite imagery and ground truth source), the architecture of the model used, how we trained the model and calibrated its predictions after training, and finally the validation exercises using external survey data.

2.1 Ground-Truth: Afrobarometer Surveys & the Sahel Desert

We use the Afrobarometer Round 6 survey to identify communities with data on infrastructure access. Afrobarometer is an Africa-wide effort to collect household-level data every 4-5 years. Round 6, which we focus on for training, was collected between 2014 and 2015, and contains information from 36 African countries. Data on infrastructure access is recorded by the enumerator at the Enumeration Area (EA) level, which contains on average 8 households. The dataset we use contains 7,022 enumeration areas, with 150 to 300 EAs per country. We exploit the fact that the Afrobarometer data was geocoded by BenYishay et al. (2017) to match each EA to the corresponding satellite images.[3] We match the centroid of the EA to the centroid of the satellite image.

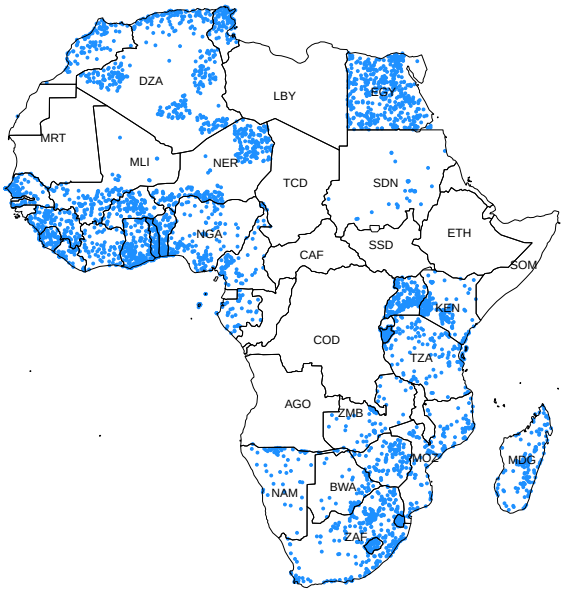
As Afrobarometer surveys, by construction, only collect data from populated areas, we do not have ground truth for locations where no one lives. To map infrastructure over the entire continent, we however need the model to learn to recognize areas where there are no households. Indeed, if the model did not learn that some places are not suitable to live and therefore, no infrastructure is built there, the risk of false positives could be high. To avoid this problem, we add to the training dataset 696 uninhabited patches from the Sahara Desert with "zero" labels for all infrastructure measures.

Figure 1 displays the location of all enumeration areas in our training dataset. We can observe that some countries have great coverage, while others have very little or no coverage at all. This highlights the need to use alternative data sources, namely satellite imagery combined

[2]See for example Albert et al. (2017); Bragilevsky and Bajić (2017); Jean et al. (2016); Oshri et al. (2018); Saavedra and Romero (2021); Xie et al. (2016); Yeh et al. (2020)

[3]The survey was geocoded ex-post, so some EAs do not have accurate coordinates. In particular, 3,053 EAs share geocoordinates with another EA. We drop these enumeration areas to focus on those for which the geolocation is accurate (N=3,969 EAs)

Figure 1: Location of Afrobarometer EAs in Africa



with machine learning techniques, rather than to do simple extrapolations from the sparse ground truth data.

2.2 Satellite Imagery

We use multispectral imagery from the Landsat 7 and 8 Surface Reflectance Tier 1 Collection and from the VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1. Landstat images are available at a 30 meter resolution, and we use 6 of its bands: 3 RGB bands and 3 infrared bands. We use the quality assessment band to remove pixels that are marked as having cloud, cloud shadow and snow in every image, and then take a median composite of every cloud-free pixel available between 2014 and 2015 for 6.72km by 6.72km patches around the centroid of the ground-truth data points (Afrobarometer enumeration areas and desert patches) described in section 2.1.

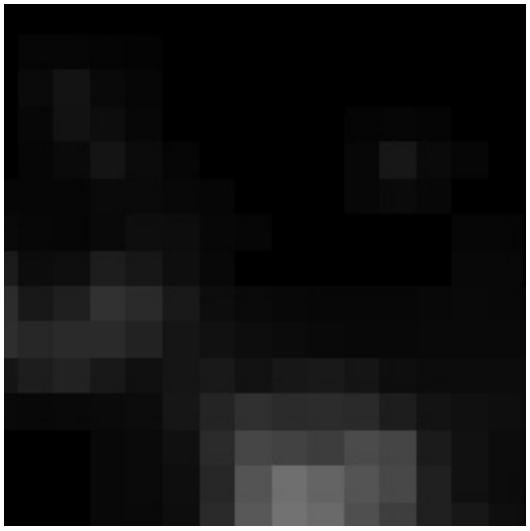
The VIIRS Nighttime Light (NTL) imagery has resolution of 15 arc seconds, approximately 500 meters per pixel, and it only contains a single band recording the radiance at night. The data product for NTL data throughout this paper has been preprocessed to filter out clouds and correct for stray light. We generate 2014-2015 median composites based on monthly data (because of the possible lack of good quality data coverage for some months). The nightlight

band is then reprojected and concatenated to the landsat images as the seventh band.[4] When we are training the model, the input imagery has size 224 pixels by 224 pixels (6.72km by 6.72km) and 7 bands, a dimension of $224 \times 224 \times 7$.

In order to illustrate how our imagery generally looks, subpanel (a) of Figure 2 displays one example of a Landsat 8 image taken from Kenya, where we only visualize the 3 RGB bands. Subpanel (b) displays an image of the same geography using VIIRS Nighttime Light imagery.

Figure 2: Satellite imagery examples

(a) LANDSAT 8 image from Kenya



(b) VIIRS image from Kenya

2.3 Machine Learning Model

We build on the approach we previously developed in Oshri et al. (2018). The first task consists of solving a multi-label binary classification problem: Using satellite images as inputs, we want to predict whether EAs have access to infrastructure or not. We train a Residual Network (ResNet), a state-of-art CNN that has been shown to have good performance for binary classification problems (He et al., 2016). The ResNet takes six bands for the Landsat data and 1 band for the VIIRS Nighttime data as input and generates a score, and makes a binary classification based on the score, while minimizing the loss function between predictions and ground truths. More technical details on the structure of the Residual Network, the weights

[4]VIIRS Nightlight Light data has different resolution from Landsat so we re-projected it to 30-meters.

initialization procedures, data processing, regularization and optimization are in Appendix C.1.

CNNs tend to overfit and generate scores very close to 0 or 1 for each observation so that they can get a small loss function. To avoid this problem, we use a post-estimation calibration technique called "temperature scaling", which ensures that the score represents the "confidence", i.e, the *likelihood* that a given area has access to specific amenities Guo et al. (2017).[5].

2.3.1 Model Performance

Out of 4,665 observations, we use 80% to train our model and 20% to test its performance, which means that we use 3,738 observations to train our model, and 927 observations to assess the performance of our final model.

Table 1 reports on the performance of the model. For each infrastructure considered, we classify a patch as 0 (no access) if its predicted probability is below 0.5, and as 1 (access) if the predicted probability is above 0.5. In the first column of Table 1, we report the mean of the ground truth outcomes in our test set: Around 55% of patches have access to electricity, 52% to piped water, and 28% to sewerage. The second column of Table 1 displays the accuracy of the model over the test set, where we calculate the number of correct predictions divided by the sample size. We can see that the model predictions are correct between 77% and 84% of the time, depending on the infastructure considered.

Next, we report *recall*—the number of correct positive predictions (when we predict there is access to infrastructure) divided by the number of actual positive observations, and *specificity*—the number of correct negative predictions (when we predict there is no access to infrastructure) divided by the number of actual negative observations. Recall ranges between 68% and 82%, and specificity is between 74% and 91%, which shows that the model performs well at predicting both instances of access and lack of access to infrastructure.[6]. Finally, we report AUCROC scores, which represent the area under the ROC curve. These scores show how well a classifier

[5]See Appendix C.2 for more details

[6]Breaking down the performance of by ground truth class (0 or 1) is important because of the class imbalance in some outcomes like sewerage. For instance, we could predict that no household has access to sewerage, and the accuracy of that prediction would be 72.4%. By looking at recall and specificity we can see that the high accuracy levels come from good predictions for both classes, and it is not being affected by class imbalances.

can distinguish between 2 classes, i.e., predicting 0 when the ground truth is 0 and predicting 1 when the ground truth is 1. This metric takes values ranging between 0 and 1, with 1 representing a perfect classifier. The area under the ROC curve ranges between 77% and 83% depending on the infrastructure considered.

Appendix Table B.1 further analyses the performance of the models separately for patches in rural vs. urban areas. Urban enumeration areas are more likely to have access to electricity, piped water, and sewerage than rural enumeration areas. It is therefore important to make sure the model performance is not simply driven by the model's ability to separate urban from rural areas. Reassuringly, Table B.1 shows that the model performs well in both rural and urban areas. As expected, recall is slightly higher in urban areas, and specificity is higher for rural areas: the model is better at identifying when there is access to infrastructure in urban areas, and at identifying when rural areas lack access to basic infrastructure.

Our model's performance is on par with similar models in this literature: Oshri et al. (2018) 's model presents accuracy levels between 67.3% and 83.2% when predicting our same three infrastructure outcome, achieving the same accuracy when predicting access to electricity, but lower accuracy when predicting access to sewerage and piped water.

Table 1: Model Performance for Infrastructure Outcomes

	Mean	Accuracy	Recall	Specificity	AUCROC
Electricity	0.547	0.832	0.822	0.843	0.833
Piped Water	0.519	0.771	0.786	0.756	0.771
Sewerage	0.276	0.847	0.684	0.909	0.796

Note: All performance metrics were computed in our test sample.

2.3.2 External validation with other datasets

We test whether our predicted measures are valid using the ground truth from new datasets with precise geolocations. We export satellite images at the corresponding locations with the same procedures as before, feed the images into our model, and generate predictions. Then, we compare our predictions with the ground truth data to check whether the model performs well in other settings. We also compare the performance of our model in these datasets against a "naive model" using nightlights only. The NL model will predict that an EA has access to

electricity, piped water, and sewerage if there is any lit patch, and no access to all infrastructure of the area is completely dark.

Our validation datasets are divided into two categories: datasets from the same time period and a different period. As the model is built to make predictions from 2014-2015 imagery, it is possible that it is particularly accurate in predicting infrastructure outcomes in the same period. We use four contemporaneous datasets at various levels (household, community) to test for within-period validity. We also test how well the model performs in earlier or later data. Specifically, for earlier periods, we use Afrobarometer Round 5, collected in 2011-2013, matched with Landast images from that same period. For later periods, we use the second round of the Greater Addis Ababa survey collected by Stanford's AUDRI initiative in 2018, paired with imagery from 2018. Details on the datasets are in Table B.2.

We present the validation results in Table B.4. We see that proposed model always significantly outperforms the NL model in predicting piped water and sewerage access. In term of predicting electricity, our model outperforms almost always, especially in rural regions. Our model performs overall better than the nighttime lights benchmark when we validate against the Afrobarometer Round 5. In particular, it does substantially better for patches in rural areas and near national borders, though it performs worse in urban areas. In the validation dataset of high schools in Ghana, nighttime lights perform slightly better when predicting access to electricity, but our model performs considerably better when predicting access to piped water. In Addis Ababa surveys, the naive nighttime lights model performs slightly better when predicting electricity only in Addis Ababa, but it performs significantly worse in piped water and sewerage. In Abidjan and rural Malawi, our model outperforms in every outcome.

2.4 A new dataset to study African development

After training, calibrating and validating the machine learning model, we can finally use it to predict infrastructure access levels across all Africa. To do this, we overlay a grid over a map of Africa, splitting the continent into 747,165 squared patches with a size of 6.72 kilometers by 6.72 kilometers. After producing our grid, we retrieve satellite imagery for every patch in the grid from the same Landsat 7 and 8 Surface Reflectance and VIIRS Nighttime Light collections

we used to train our model. We input these images to our calibrated machine learning model and can then predict, for any given patch, the probabilities that it has access to piped water, sewage systems and electricity.

We complement our dataset of predicted outcomes with readily available rasters of data containing information about development related outcomes, re-projecting them to match our grid size or taking the average of the pixels of these rasters for each patch of our grid.[7] In particular, we use VIIRS Nighttime Light for 2014 and 2015 to generate the average nighttime luminosity for each patch in the grid; we compute average population density data for the year 2015 using Data For Good at Meta's High Resolution Population Density Maps (Tiecke et al., 2017); and we use the GRIP global roads database to compute the distance to the closest road for each patch in the grid (Meijer et al., 2018).

We incorporate relevant geographic control variables into the dataset: For each patch of land we compute the distance to the capital and the distance to the coast using detailed maps of current national borders in Africa (GADM, 2018). We also compute the distance to the closest petroleum and diamond sources inside and outside the country, using the PETRODATA v1.2 dataset from the Peace Research Institute Oslo (Lujala et al., 2007) and the DIADATA v1a dataset from the same institution (Gilmore et al., 2005). We obtained data about Malaria Suitability (P. Falciparum suitability) for 2010 from (Gething et al., 2011). We retrieved a measure of land suitability for agriculture from The Atlas of the Biosphere (Ramankutty et al., 2002). Finally, we incorporate information about the elevation and the distance from the equator for each patch of land in our dataset, and an indicator variable for landlocked countries.

Table 2 presents descriptive statistics from the resulting dataset. Patches on the African continent have on average a 29% probability of having access to electricity. The corresponding figures for sewerage and piped water are 14% and 29%, respectively. There is substantial dispersion. In contrast, nighttime lights intensity, expressed as the average radiance of a patch, is highly skewed: only around 10% of patches have a positive radiance value at night.

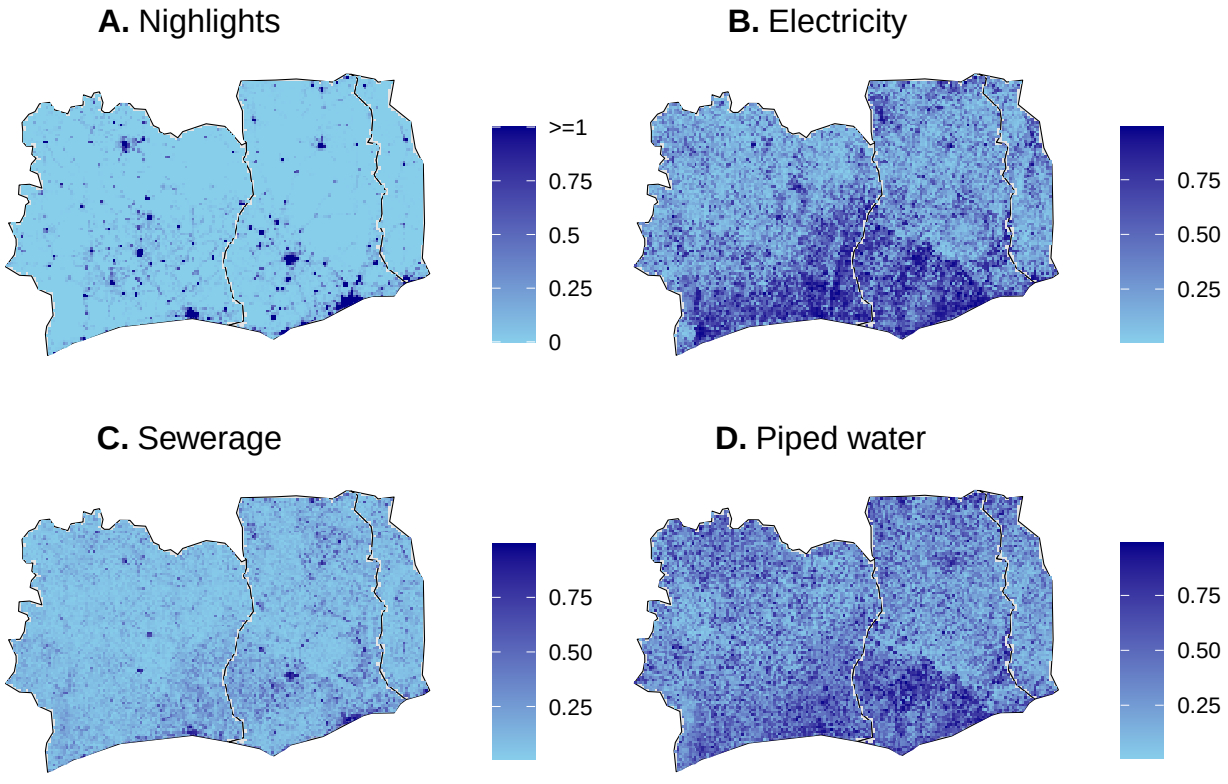
[7] see Appendix D for more details.

Table 2: Summary Statistics: Patch-Level Outcomes

Variable	Mean	Std. Dev	Min	P25	Median	P75	Max	Obs
Electricity	0.29	0.26	0.00	0.08	0.18	0.45	1.00	746,475
Sewerage	0.14	0.14	0.01	0.07	0.10	0.16	1.00	746,475
Water	0.29	0.22	0.00	0.11	0.23	0.44	0.99	746,475
Nightlights	0.10	1.61	0.00	0.00	0.00	0.00	427.53	746,605
Distance to the closest road	17.05	35.63	0.00	0.00	3.41	18.56	3,625.39	731,145
Population Density	33.19	278.19	0.00	0.00	0.00	7.35	48,962.05	747,158

To illustrate the value added of our new dataset, Figure 3 zooms in on a specific region (the three West African countries of Côte d'Ivoire, Ghana and Togo). While the night lights map would suggest only urban centers have access to electricity, our model predictions reveal substantial variation across patches within rural areas.

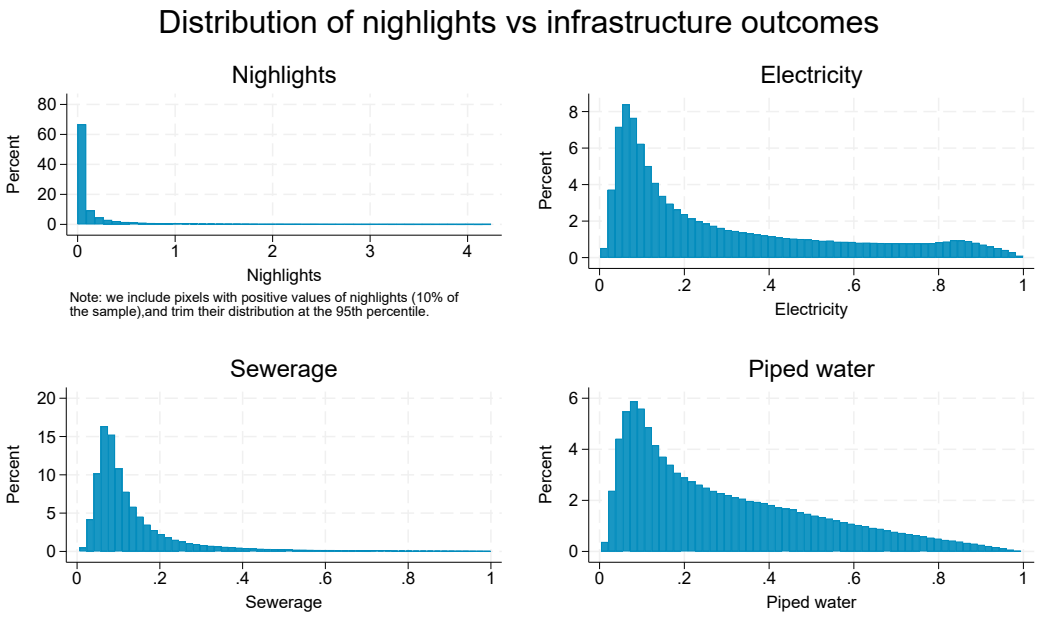
Figure 3: Predictions: Zooming in on Cote d'Ivoire, Ghana and Togo



This pattern is not specific to the West Africa area shown in Figure 3. Figure 4 plots the distribution of our predicted infrastructure outcomes against the distribution of nightlights *for*

the subset of patches with non-zero night lights. Even within this selected sub-sample, there is very little variation in the night lights information, with 80% of patches having extremely low radiance. In contrast, for this same sub-sample, our electricity and other infrastructure predictions exhibit rich variation.

Figure 4: Comparison of outcome distributions



Because it was so far the only available outcome, most of the previous literature on comparative development has focused on night lights density (Chen and Nordhaus, 2011; Henderson et al., 2012; Michalopoulos and Papaioannou, 2013, 2014; Moscona et al., 2020; Canning et al., 2022). While this has enabled the literature to make great strides, we believe that it is now time to go beyond night time lights. We illustrate this in the next two sections, in which we use our newly compiled dataset to revisit some old debates in the literature on the determinants of local development.

3 Application I: Institutions and Development

While there is a robust correlation between institutional quality and growth in cross-country analyses, disentangling the pathways of influence has proved elusive due to a lack of exogenous variation in institutional quality. Focusing on African development, Michalopoulos

and Papaioannou (2014) exploited the arbitrary drawing of borders during the colonial period to estimate the impact of national institutions on economic development using a regression discontinuity approach. The only outcome available at the patch level at the time was night lights density. Their analysis yielded a (surprising) null result: national institutions do not impact development levels, in sharp contrast with the earlier findings of Acemoglu et al. (2001). In this section, we revisit this question, bringing to bear the comprehensive and rich dataset we just described. We find that national institutions do matter for infrastructure development, though the magnitude of the effect is much more modest than the "development" effect estimated in Acemoglu et al. (2001).

3.1 Historical Background

The borders that divide African countries are largely the result of boundary-drawing practices from the colonial era, which mostly ignored the sociological, cultural, and historical aspects of the peoples they separated (McCauley and Posner, 2015). These colonial political borders date from the Berlin Conference of 1884-85, during which France, Germany, Great Britain, Italy, Portugal and Spain partitioned the African continent. These final borders did not always followed geographical features: 30% of Africa's boundaries are straight lines according to Boggs (1940), and 44% follow astronomical lines according to Barbour (1961). McCauley and Posner (2015) estimates that around 80% of borders follow meridian, parallels or other rectilinear or curved lines. As Griffiths (1986) notes, European colonizers laid borders with very odd shapes and varied sizes: They created very long narrow states, like The Gambia and Togo, where development concentrated in very short sea coasts that are very isolated from the rest of the country; perhaps most problematic, the drawn borders caused 14 countries to be landlocked, more than there are in the rest of the world combined. These arbitrary colonial boundaries mostly survived independence processes and still separate today's modern African countries.

Several ethnic groups inhabited the African continent before colonization, and the limits of their territories were very different from the boundaries erected by colonial powers. Asiwaju (1985) identifies 177 "culture areas", and every country border in Africa cuts through at least one. For instance, the boundaries of Burkina Faso cut through 21 culture areas. One of the most

comprehensive atlas of the ethnic cultures inhabiting Africa before its colonization is arguably Murdock (1959), who compiled anthropological descriptions of most African ethnicities in the mid to late nineteenth century, describing the geography of the territories they inhabited, their languages, economies and how their societies were organized. Murdock produced a tribal map of Africa to accompany his book; this map, which some historians have deemed possibly quite approximative, was first digitized by Nunn (2008).

3.2 National institutions and development

Following Michalopoulos and Papaioannou (2014), we intersect the "Tribal Map" from Murdock (1959) with the map of modern national borders in Africa (GADM, 2018) to identify ethnic groups that were partitioned by national borders, and use a spatial regression discontinuity design, exploiting the arbitrary placement of national borders within the African continent by European colonizers. The idea behind this is to compare the development levels of areas near a national border—areas that share similar geographic features and that were exposed to the same pre-colonial institutions and customs, but that are exposed to different current national institutions. Some ethnic groups were split into 3 or more countries—in such cases, we focus on the pair of countries with the biggest combined ethnic area. There are a total of 243 partitioned ethnicities across 47 countries. In our dataset, 318,887 patches belong to one of these partitioned ethnicities.

We estimate the following equation:

$$\text{Dev}_{p,e,c} = \alpha + \beta\, \text{National Institutions}_{e,c}^{\text{high}} + f(BD_{p,e,c}) + X'_{p,e,c}\Gamma + \delta_e + \varepsilon_{p,e,c} \tag{1}$$

Where $\text{dev}_{p,e,c}$ is one of our development outcomes for patch p in the ethnicity e in country c, $\text{National Institutions}_{e,c}^{\text{high}}$ is a dummy denoting if country c has a higher institutional quality than the country on the other side of the border, $f(BD_{p,e,c})$ is a first degree polynomial (estimated separately on each side of the border) in the running variable, $BD_{p,e,c}$, the distance from the centroid of patch p to the relevant national border that partitions ethnicity e, $X'_{p,e,c}$ is a vector of characteristics and δ_e is a ethnicity level fixed effect. Standard errors are clustered at the country level, the level at which the independent variable of interest (national institutions)

varies. We follow Calonico et al. (2014) to estimate the optimal RD bandwidth.[8]

Our primary development outcomes are the predicted infrastructure access measures for the year 2015 described in Subsection 2.4, along with an indicator variable indicating if a patch of land is lit according to Night Lights 2015, the logarithm of Population Density 2015, the distance to the closest road up to the tertiary level for the patch of land and an index of our outcomes derived from the first component of a Principal Components Analysis. We control for distance to the capital, distance to the coast, closest diamond and petroleum sources inside and outside the country, temperature suitability for Malaria, land suitability for agriculture, elevation, distance from the equator and whether a country is landlocked.

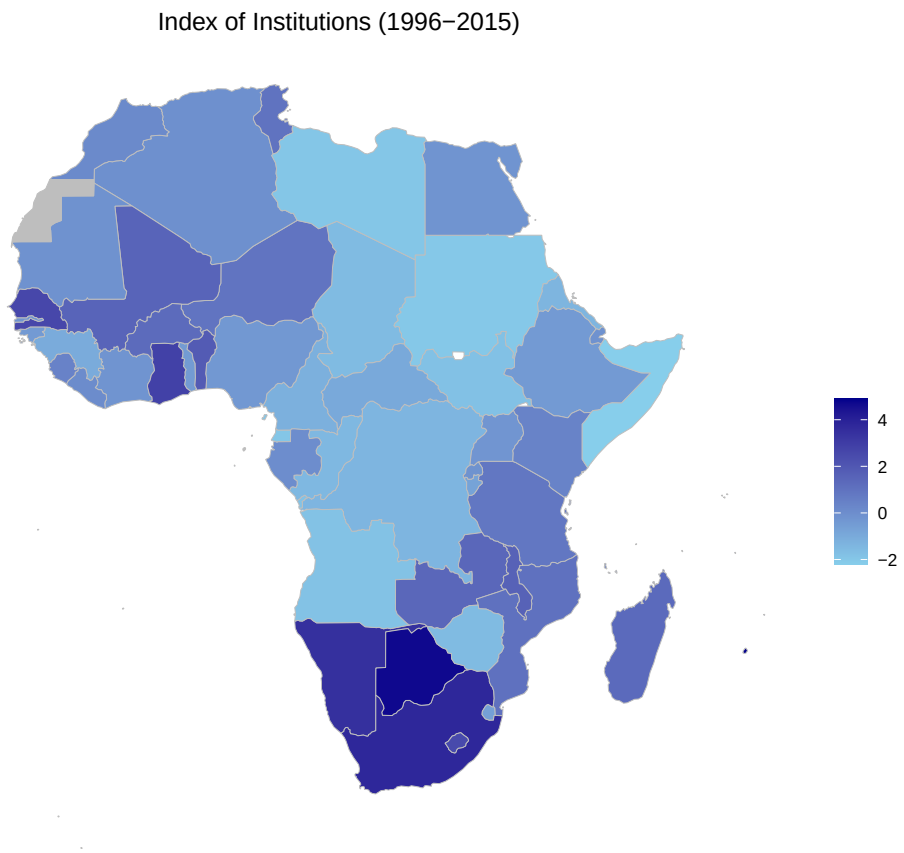
As for the independent variable (institutional quality), there are several measures available in the literature, and among them, we use 4 measures: Rule of Law and Control of Corruption from the Worldwide Governance Indicators (Kaufmann et al., 2011), the Polity IV score from the Center for Systemic Peace (Marshall et al., 2019) and the Electoral Democracy index from Varieties of Democracy Dataset version 11 (Coppedge et al., 2021). Rule of Law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Control of Corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests. The Polity score of a country captures this regime authority spectrum on a 21-point scale ranging from -10 (hereditary monarchy) to +10 (consolidated democracy). The Electoral Democracy index captures to what extent the ideal of electoral democracy in its fullest sense is achieved, with politicians that are responsive to citizens, competitive and fair elections, and freedom of expression and press.

For each institutional variable we use the average value over the period 1996-2015. We also combine them in an index of institutions, using the first component of a Principal Component Analysis. There is heterogeneity in these various measures across African nations, as shown in

[8]To use the same bandwidth for all outcomes, we create a PCA index of the key development outcomes and use it as the dependent variable in a uniform kernel. Doing this, we obtained an optimal bandwidth of 49.5 kilometers for each side of the border.

Figure 5.

Figure 5: National Institutions: Index of institutions



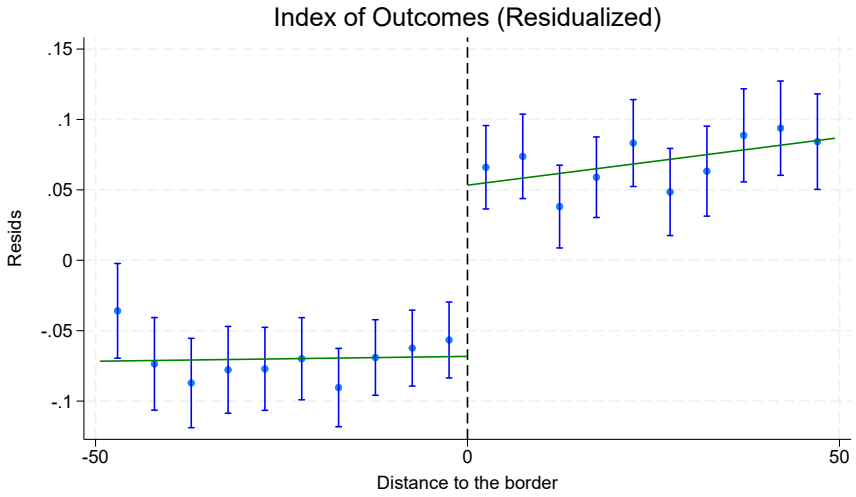
3.2.1 National institutions and development: Results

Figure 6 plots the results of the analysis of how overall institution quality matters for overall development. We see a clear jump at the border, with patches in the country with higher institutional quality exhibiting higher levels of infrastructure development. We show the breakdown by outcome in Figure A.1. There is a clear discontinuity for four of six outcomes.[9].

Table 3 shows the coefficient estimates and standard errors. In Panel A we report the results for the PCA index of institutions, while in Panel B we show the results for each of the four measures of institutional quality considered separately. We see very consistent, positive and

[9]We show the non-residualized plots in Figure A.2 and Figure A.3

Figure 6: Discontinuity around national borders



Note: The country with the better institutions is on the right (positive distance to the border).

significant results across outcome measures and across institutional quality measures. Crossing from a country with lower institutional quality to a country with higher institutions quality increases the overall infrastructure access index by 0.19 standard deviations, the likelihood of a patch being lit increases by 3.54 percentage points, the likelihood of having access to electricity, sewerage and piped water increases by 1.72, 1.12 and 1 percentage points, respectively. Patches in countries with higher institutional quality are also on average 4.73 kilometers closer to the closest road.

Interpreting the magnitudes To get a sense of what it means to "cross" a national border to the side with better institutions, Figure 7 reports the distribution of the differences in our index of institutions for all country pairs in our sample. The mean gap is 1.07, which corresponds to 0.77 standard deviations in the country distribution of the institutions' quality index. As an example, crossing from the Democratic Republic of Congo to Uganda, the quality of institutions index increases by 1.1 points, reflecting Uganda's higher Rule of Law and Control of Corruption scores. When crossing from Zimbabwe to Mozambique we see an increase in the index of institutions of 2.5 points, due to Mozambique's much higher scores in Rule of Law, Control of Corruption and Polity IV. Finally, going from Côte d'Ivoire (which experienced two civil wars since the beginning of the 21st century) to Ghana corresponds to a jump in the

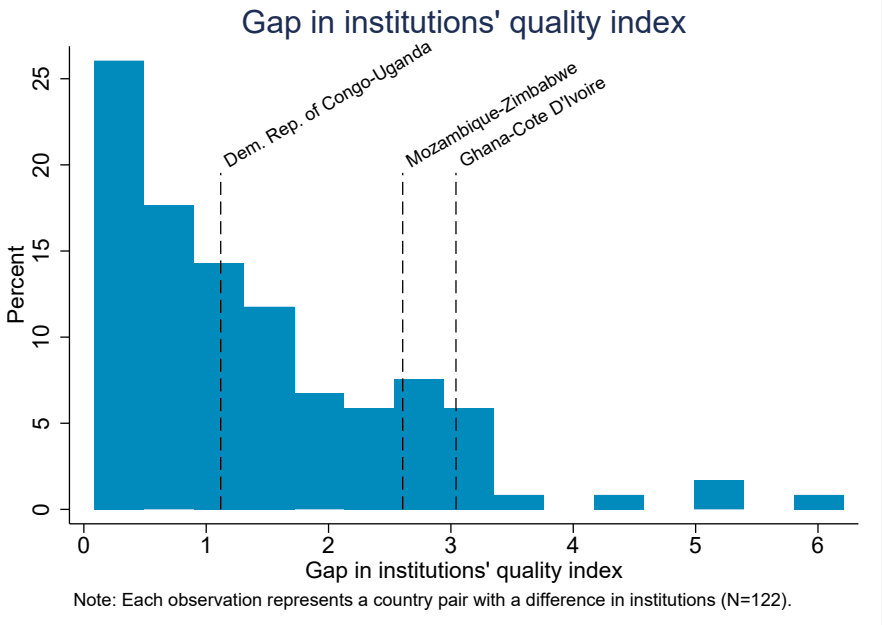
Table 3: Institutions and Development: RD Results

Outcome	(1) Index	(2) Nightlights	(3) Electricity	(4) Sewerage	(5) Piped water	(6) Dist. Road	(7) Pop. Density
Panel A: index of institutional quality (PCA)							
Index of institutions (PCA)	0.1949***	0.0354***	0.0172**	0.0119**	0.0105**	-4.7360*	0.2756
	(0.054)	(0.011)	(0.008)	(0.005)	(0.005)	(2.716)	(0.221)
	{0.00}	{0.00}	{0.03}	{0.02}	{0.04}	{0.09}	{0.22}
Observations	82,082	82,088	82,082	82,082	82,082	82,088	82,088
Panel B: breakdown by measure of institutional quality							
Rule of Law	0.0846*	0.0362***	0.0069	0.0085**	0.0097**	6.1918**	0.1619
	(0.046)	(0.009)	(0.007)	(0.004)	(0.004)	(2.477)	(0.210)
	{0.07}	{0.00}	{0.31}	{0.04}	{0.04}	{0.02}	{0.45}
Control of Corruption	0.1019**	0.0301***	0.0063	0.0075**	0.0081*	-0.9470	0.0487
	(0.048)	(0.009)	(0.006)	(0.004)	(0.004)	(2.324)	(0.198)
	{0.04}	{0.00}	{0.32}	{0.04}	{0.07}	{0.69}	{0.81}
Polity IV Score	0.1385**	0.0330***	0.0172**	0.0065*	0.0071	-6.1547**	-0.0929
	(0.060)	(0.012)	(0.007)	(0.004)	(0.005)	(2.438)	(0.265)
	{0.03}	{0.01}	{0.02}	{0.08}	{0.12}	{0.02}	{0.73}
Electoral democracy	0.2147***	0.0418***	0.0174**	0.0065	0.0098*	-4.4450	0.5798***
	(0.052)	(0.012)	(0.008)	(0.004)	(0.005)	(2.663)	(0.175)
	{0.00}	{0.00}	{0.03}	{0.12}	{0.06}	{0.10}	{0.00}
Observations	82,082	82,088	82,082	82,082	82,082	82,088	82,088

Notes: We include ethnicity level fixed effects, standard errors are clustered at the country level and are reported in parentheses, p-values are reported in curly brackets.

institutional quality index of more than 3 points.

Figure 7: Gaps in index of institutions across national borders



Comparing the coefficient estimates in Table 3 to the interquartile ranges shown in Table B.5, we see that crossing a national border to the side with better institutions increases access to

infrastructure by up to 1/3 of the interquartile range of the country-level distribution of that infrastructure outcome across Africa. Rescaling, this means a one standard deviation gain in institutional quality increases access to infrastructure by at most 0.33/0.77=42%. This is not a trivial effect, but it is much smaller than the estimates in Acemoglu et al. (2001), which find that a one standard deviation increase in protection against expropriation risk increases GDP per worker by between 100% and 300%, suggesting that the results in Acemoglu et al. (2001) may have been upward bias—in line with the conclusion of Pande and Udry (2005), which show that the Acemoglu et al. (2001) approach would imply a one standard deviation gain in institutional quality would (implausibly) move a country from the 25th percentile to the 75th percentile in the headcount ratio distribution.

3.2.2 Robustness checks

Table 4 shows a battery of robustness checks. We focus on the the specification using the PCA index of institutions. Panel A reproduces the results from Panel A of Table 3 for benchmarking. Panel B shows the results excluding controls. The results are unchanged. Panel C focuses on borders with large discontinuities (defined as a difference in institutional quality across the two sides of the border above the median). Unsurprisingly, the coefficients increase, consistent with a "dose response" effect of institutional quality. Panel D focuses on the subsample of borders with the same colonizer on both sides.[10] The sample size shrinks, but most of the results stand. Finally, Panel E shows that the results mostly hold when we focus on patches far away from the capital of their respective countries.[11]

Figure 8 checks whether the results are affected by the RD specification choices. Panel A shows that the estimated impact of institutional quality on outcomes is unaffected by the polynomial degree choice. Panel B shows that the effect, if anything, increases with bandwidth size. Finally,

[10] In appendix F we present information on the colonizer both (a) following the "scramble for Africa" and (b) at the time of independence, and present results using one or the other.

[11] A straightforward way to define patches away from capital cities would be to just exclude patches within a certain radius of their respective capitals. A potential problem with this is that it could lead to ethnicities with a very unbalanced number of patches between the two sides of the border, if only one side is close to the capital of a country. To solve this, for each ethnicity we assign each patch a "match" on the other side of the border (its symmetric), so each patch has a unique counterpart on the other side of the border that is at a very similar distance from the border. To match them, we use the Hungarian Algorithm. We then focus on the sample of paired patches and estimate the previous specification only for pairs of patches where neither patch is within 300 km of the capital of their respective country.

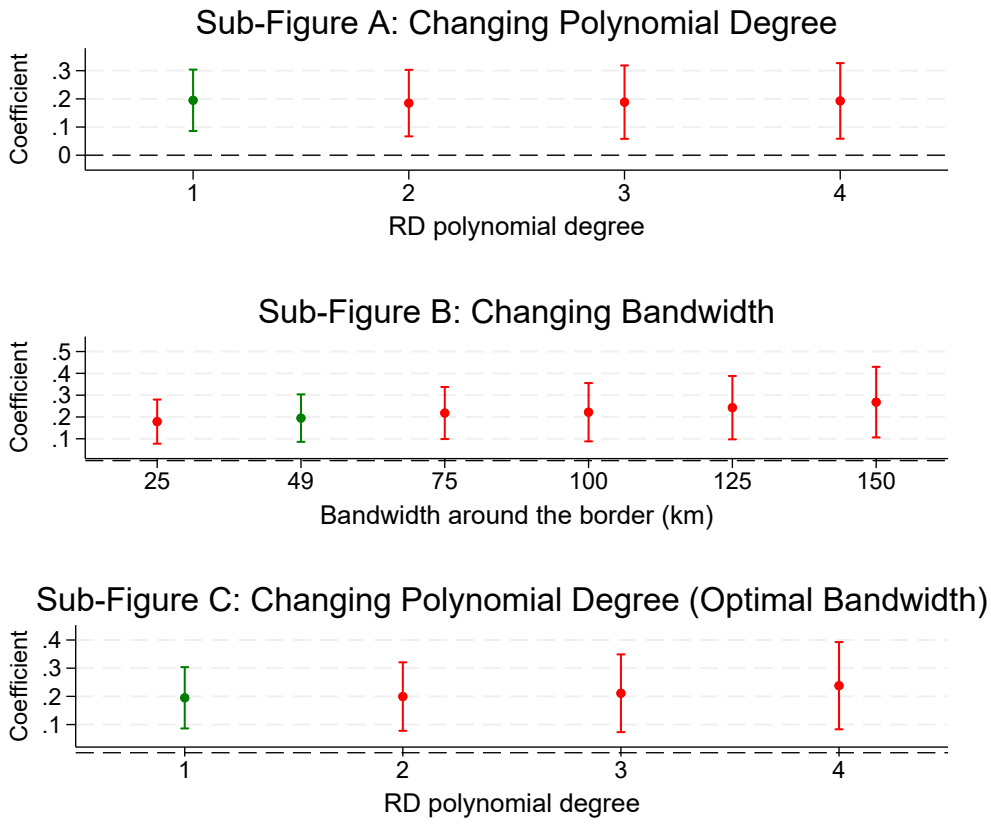
Table 4: Alternative RD specifications

Outcome	Index	Nightlights	Electricity	Sewerage	Piped water	Dist. Road	Pop. Density
Panel A. Main specification:							
Index of institutions (PCA)	0.1949***	0.0354***	0.0172**	0.0119**	0.0105**	-4.7360*	0.2756
	(0.054)	(0.011)	(0.008)	(0.005)	(0.005)	(2.716)	(0.221)
	{0.00}	{0.00}	{0.03}	{0.02}	{0.04}	{0.09}	{0.22}
Observations	82,082	82,088	82,082	82,082	82,082	82,088	82,088
Panel B. No controls:							
Index of institutions (PCA)	0.1650***	0.0284***	0.0111**	0.0068**	0.0075*	-2.3909	0.5415**
	(0.051)	(0.010)	(0.005)	(0.003)	(0.004)	(2.707)	(0.224)
	{0.00}	{0.01}	{0.03}	{0.02}	{0.05}	{0.38}	{0.02}
Observations	82,083	82,089	82,083	82,083	82,083	82,089	82,089
Panel C. Large discontinuities:							
Index of institutions (PCA)	0.2190**	0.0598***	0.0124	0.0181**	0.0104	-2.5703	0.2375
	(0.083)	(0.017)	(0.012)	(0.007)	(0.009)	(2.078)	(0.217)
	{0.01}	{0.00}	{0.32}	{0.01}	{0.24}	{0.22}	{0.28}
Observations	43,184	43,190	43,184	43,184	43,184	43,190	43,190
Panel D. Ethnicities with the same colonizer:							
Index of institutions (PCA)	0.1312*	0.0467**	0.0317***	0.0182**	0.0077	-0.5426	-0.6146***
	(0.065)	(0.020)	(0.010)	(0.008)	(0.009)	(4.022)	(0.212)
	{0.05}	{0.02}	{0.00}	{0.03}	{0.40}	{0.89}	{0.01}
Observations	29,667	29,673	29,667	29,667	29,667	29,673	29,673
Panel E. Neither patch within 300 km of the capital:							
Index of institutions (PCA)	0.1733***	0.0286***	0.0105	0.0072*	0.0092	-7.7710**	0.2536
	(0.056)	(0.010)	(0.007)	(0.004)	(0.006)	(3.374)	(0.249)
	{0.00}	{0.01}	{0.12}	{0.06}	{0.15}	{0.03}	{0.32}
Observations	44,397	44,403	44,397	44,397	44,397	44,403	44,403

Notes: We include ethnicity level fixed effects, standard errors are clustered at the country level and are reported in parentheses, p-values are reported in curly brackets.

in Panel C we vary the polynomial degree and re-compute the optimal bandwidth (our optimal bandwidth is increasing in the polynomial degree, as reported in Table B.6). The magnitude of the coefficient is very stable.[12]

Figure 8: Robustness of results to RD polynomial degree and bandwidth



Another possible concern is that we are considering six outcomes, so at a 5% significance level we have a 26% probability ($1 - 0.95^6$) of a false rejection for each one of our institutional measures. In Table B.7 we show that adjusting the p-values to account for multiple hypotheses testing does not alter the key finding that institutional quality matters for infrastructure access. Specifically, for each institutional variable we compute the sharpened False Discovery Rate p-values described in Anderson (2008), using the main RD specification with fist degree RD polynomials, the optimal bandwidth, controls, fixed effects and clustered standard errors. For each specification shown in Table 3, we report the regression coefficients, p-values computed without clustering, clustered p-values, and sharpened False Discovery Rate

[12]These robustness checks are shown for each outcome separately in Figure A.4, Figure A.6 and Figure A.7.

clustered p-values.

4 Application II: Distributive Politics

Distributive politics, or how politicians and policy makers allocate public goods and services, is a major determinant of welfare, specially when it comes to infrastructure, because connecting an area to the electrical grid, the sewage system, improving access to piped water, or building roads to improve their connectivity can also lead to further opportunities for economic development. Starting with Bates (1974), ethnicity has been the focus of analyses of distributive allocations in Sub-Saharan Africa: see for example Burgess et al. (2015) and more recently Dreher et al. (2019). This literature typically focuses on one type of good in one country. But, as shown by Kramon and Posner (2013), there are many goods and services allocated simultaneously by governments, and if one group receives more roads, another may receive more health care. Golden and Min (2013) thus call for studies that consider multiple goods before clear patterns of political favoritism can be established.

We use our dataset to study the extent to which the allocation of infrastructure across geographical units within a country is consistent with ethnic favoritism, a phenomenon where politicians prioritize the geographies where their ethnicity resides when allocating public resources. We compute average infrastructure access at the second-level administrative unit (the next smallest unit after a state or region), and merge the dataset with information on the ethnicity of presidents since independence (or 1960 for countries that were never colonized)[13]. We estimate the following equation:

$$\text{Dev}_{g,c} = \alpha + \beta\, \text{Ties}_{g,c} + X'_{g,c}\Psi + \delta_c + \varepsilon_{g,c} \tag{2}$$

Where $\text{Dev}_{g,c}$ represents the average development outcome at second-level administrative unit g in country c, $\text{Ties}_{g,c}$ is an indicator of geographical unit g containing the birthplace of any president ever elected in country c (a proxy for their ethnic homeland), which means one of these presidents has political ties to unit g, $X'_{g,c}$ is a vector of geographic controls, and δ_c are country fixed effects. Standard errors are clustered at the country level.

[13] We use the dataset built by Dreher et al. (2019) covering the period between 2000 to 2012, and add hand-collected data between the years 1960 to 2000.

In Panel A of Table 5 reports the result of estimating Equation (2), where we find a strong positive correlation between political ties and infrastructure access across all infrastructure measures considered. In Panel B we control for areas being far from the capital and we interact it with our political tie variable. Our findings suggest areas political ties affect more the allocation of infrastructure in areas closer to national capitals compared to our baseline, geographies that are close to the capital but have no political ties. Areas that are far from the capital and have no political ties have less access to infrastructure, but areas that are far from the capital that have ties to a president don't experience reduced access to infrastructure (their net effect is not statistically significant for any outcome). This means that political ties prevent geographies from suffering the negative effects of being far away from capital cities.

Table 5: Allocation of basic infrastructure

Outcome	**Index**	**Nightlights**	**Electricity**	**Sewerage**	**Piped water**	**Dist. Road**
Panel A:						
Political tie	1.1637***	0.8428***	0.0990***	0.1209***	0.1081***	-0.4414**
	(0.238)	(0.267)	(0.021)	(0.032)	(0.024)	(0.210)
	{0.00}	{0.00}	{0.00}	{0.00}	{0.00}	{0.04}
Outcome mean	1.88	0.48	0.46	0.23	0.44	1.35
Observations	4,633	4,633	4,633	4,633	4,633	4,633
$\boldsymbol{R^2}$	0.58	0.59	0.53	0.64	0.48	0.27
Panel B:						
Political tie	1.2912***	1.0315***	0.0992***	0.1500***	0.1226***	-0.4522*
	(0.264)	(0.278)	(0.027)	(0.034)	(0.029)	(0.263)
	{0.00}	{0.00}	{0.00}	{0.00}	{0.00}	{0.09}
Far from capital	-0.4074**	-0.1127	-0.0473**	-0.0283**	-0.0370**	0.0623
	(0.159)	(0.080)	(0.018)	(0.011)	(0.017)	(0.367)
	{0.01}	{0.16}	{0.01}	{0.02}	{0.04}	{0.87}
Political tie × Far from capital	-0.6657**	-0.6915**	-0.0288	-0.1133***	-0.0701*	0.0731
	(0.322)	(0.273)	(0.038)	(0.033)	(0.041)	(0.501)
	{0.04}	{0.01}	{0.46}	{0.00}	{0.09}	{0.88}
Outcome mean	1.88	0.48	0.46	0.23	0.44	1.35
Observations	4,633	4,633	4,633	4,633	4,633	4,633
$\boldsymbol{R^2}$	0.60	0.61	0.54	0.66	0.50	0.27
p-value of sum of coefficients	0.28	0.20	0.38	0.56	0.48	0.43

Notes: The unit of observation is a subnational geographic unit. The sample administrative level-2 (ADM2) units from 39 countries and administrative level-1 units from 7 countries for which ADM2 was not available. We include country-level fixed effects, and standard errors are clustered at the country level and are reported in parentheses. We control for distance to the capital, distance to the coast, closest diamond and petroleum sources inside and outside the country, temperature suitability for Malaria, land suitability for agriculture, elevation, distance from the equator, and area of the administrative unit. ***, **, and * indicate statistical significance at 1%, 5%, and 10% significance levels, respectively.

Studying sub-national allocation of roads in Kenya, Burgess et al. (2015) find that democracy constrains ethnic favoritism: expenditures on roads are twice larger in districts that share the ethnicity of the president during non-democratic periods, but there is no difference during

periods of democracy. Building on this, Panel A of Table 6 examines whether parochialism is mitigated by good institutions. In other words, we ask: does having higher quality institutions keep presidents in check? For this, we add an interaction term between political tie and the country-level index of institutional quality described in section 3. The estimates in column 1 imply that a one standard deviation increase in institutional quality reduces the role of political ties in the allocation of infrastructure by about 20%. These mitigating effects range between 9.8% for distance to the road to 27% for electricity.

Our last analysis, shown in Panel B of Table 6, considers the role of foreign aid. We split countries into two groups, those receiving aid above the median, and those who receive aid below the median in terms of total *per capita* aid received between 2000 and 2014, to then interact our proxy of ethnic favoritism with an indicator variable of receiving high total international aid. We find not only that areas with political ties have on average better access to infrastructure, but that international aid further exacerbates this effect, which can be seen as

5 Conclusion

This paper utilizes satellite imagery and deep learning models to generate novel data on infrastructure access for every single 6.72km $\times$ 6.72km area of Africa. We combine household surveys and satellite imagery to train a machine learning model to predict infrastructure access (access to piped water, sewerage and electricity). Models achieve accuracy levels of 77.1% to 84.7%. Using different metrics to evaluate model performance, we find that recall ranges between 68% and 82%, and specificity is between 74% and 91%, which indicates that our models predict both instances of access and lack of access to infrastructure accurately. Additionally, our models perform well out of sample, and particularly dominate alternatives in remote areas.

We showcase the value of this data with a few applications, revisiting some old debates in the political economy literature. We find that the quality of national institutions over the past 20 years has played a significant role in infrastructure access, as well as its allocation across areas within a country.

Our models could be applied to imagery over time, to generate panel data on infrastructure

Table 6: Political Ties and Institutional Quality

Outcome	Index	Nightlights	Electricity	Sewerage	Piped water	Dist. Road
Panel A: Institutional quality						
Political tie	1.3746***	1.0350***	0.1144***	0.1483***	0.1230***	-0.4005
	(0.272)	(0.298)	(0.026)	(0.035)	(0.028)	(0.240)
	{0.00}	{0.00}	{0.00}	{0.00}	{0.00}	{0.10}
Political tie × Index of institutions	-0.3053**	-0.1381	-0.0309**	-0.0298	-0.0215*	0.0396
	(0.144)	(0.130)	(0.014)	(0.018)	(0.012)	(0.126)
	{0.04}	{0.30}	{0.03}	{0.10}	{0.09}	{0.75}
Outcome mean	1.88	0.48	0.46	0.23	0.44	1.35
Observations	4,633	4,633	4,633	4,633	4,633	4,633
$\boldsymbol{R^2}$	0.58	0.60	0.53	0.65	0.48	0.27
Panel B:International aid						
Political tie	0.9937***	0.6822	0.0876***	0.1066**	0.0871***	-0.4704
	(0.337)	(0.409)	(0.030)	(0.049)	(0.027)	(0.320)
	{0.01}	{0.10}	{0.01}	{0.03}	{0.00}	{0.15}
Political tie × High total aid	0.5828	0.6777	0.0320	0.0673	0.0618	0.1981
	(0.532)	(0.623)	(0.044)	(0.072)	(0.057)	(0.442)
	{0.28}	{0.28}	{0.47}	{0.36}	{0.29}	{0.66}
Outcome mean	1.88	0.48	0.46	0.23	0.44	1.35
Observations	4,633	4,633	4,633	4,633	4,633	4,633
$\boldsymbol{R^2}$	0.58	0.61	0.53	0.65	0.48	0.27
p-value of sum of coefficients	0.00	0.00	0.00	0.00	0.00	0.37

Notes: The unit of observation is a subnational geographic. The sample administrative level-2 (ADM2) units from 39 countries and administrative level-1 units from 7 countries for which ADM2 was not available. We include country fixed effects. Standard errors are clustered at the country level and are reported in parentheses. High total aid is 1 if a country's total aid (summing World Bank and Chinese aid) is above the median, 40.34 constant 2017 USD per capita. World Bank aid data is downloaded from World Bank Projects website. Chinese aid is from AidData's Geocoded Global Chinese Official Finance Dataset (Version 1.1.1). We control for distance to the capital, distance to the coast, closest diamond and petroleum sources inside and outside the country, temperature suitability for Malaria, land suitability for agriculture, elevation, distance from the equator, and area of the administrative unit. ***, **, and * indicate statistical significance at 1%, 5%, and 10% significance levels, respectively.

access. They could also be adapted to other continents, to generate a world-level dataset.

References

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

Albert, Adrian, Jasleen Kaur, and Marta C Gonzalez, "Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale," in "Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining" 2017, pp. 1357–1366.

Alsan, Marcella and Claudia Goldin, "Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920," *Journal of Political Economy*, 2019, *127* (2), 586–638.

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 2008, *103* (484), 1481–1495.

Ashraf, Nava, Edward Glaeser, Abraham Holland, and Bryce Millett Steinberg, "Water, Health and Wealth: The Impact of Piped Water Outages on Disease Prevalence and Financial Transactions in Zambia," *Economica*, 2021.

Asiwaju, Anthony Ijaola, *Partitioned Africans: Ethnic relations across Africa's international boundaries, 1884-1984*, C. Hurst & Co. Publishers, 1985.

Barbour, Kenneth M, *A geographical analysis of boundaries in inter-tropical Africa* 1961.

Bates, Robert H, "Ethnic competition and modernization in contemporary Africa," *Comparative political studies*, 1974, *6* (4), 457–484.

BenYishay, A., R. Rotberg, J. Wells, Z. Lv, S. Goodman, L. Kovacevic, and D. Runfola, "Geocoding Afrobarometer Rounds 1-6: Methodology & Data Quality," 2017.

Boggs, Samuel Whittemore, *International boundaries: a study of boundary functions and problems*, Ams PressInc, 1940.

Bragilevsky, Lior and Ivan V Bajić, "Deep learning for Amazon satellite image analysis," in "2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)" IEEE 2017, pp. 1–5.

Buchhorn, Marcel, Myroslava Lesiv, Nandin-Erdene Tsendbazar, Martin Herold, Luc Bertels, and Bruno Smets, "Copernicus global land cover layers—collection 2," *Remote Sensing*, 2020, *12* (6), 1044.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel, "The value of democracy: evidence from road building in Kenya," *American Economic Review*, 2015, *105* (6), 1817–51.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Canning, David, Marie Christelle Mabeu, and Roland Pongou, "Colonial origins and fertility: Can the market overcome history?," *Available at SSRN 4070963*, 2022.

Carroll, ML, CM DiMiceli, MR Wooten, AB Hubbard, RA Sohlberg, and JRG Townshend, "MOD44W MODIS/Terra Land Water Mask Derived from MODIS and SRTM L3 Global 250m SIN Grid V006 [Data set]," *NASA EOSDIS Land Processes DAAC*, 2017.

Center for International Earth Science Information Network, CIESIN, "Gridded Population of the World, Version 4 (GPWv4): Population Density," 2016. Accessed August 9, 2021.

Chen, Xi and William D Nordhaus, "Using luminosity data as a proxy for economic statistics," *Proceedings of the National Academy of Sciences*, 2011, *108* (21), 8589–8594.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Nazifa Alizada, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerlow, Adam Glynn, Allen Hicken, Garry Hindle, Nina Ilchenko, Joshua Krusell, Anna Luhrmann, Seraphine F. Maerz, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Johannes von Romer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundstrom, Ei tan Tzelgov, Yi ting Wang, Tore Wig, Steven Wilson, and Daniel Ziblatt., "V-Dem Country-Year/Country-Date Dataset v11.1," 2021.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Parienté, and Vincent Pons, "Happiness on tap: Piped water adoption in urban Morocco," *American Economic Journal: Economic Policy*, 2012, *4* (4), 68–99.

Dinkelman, Taryn, "The effects of rural electrification on employment: New evidence from South Africa," *American Economic Review*, 2011, *101* (7), 3078–3108.

Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley C Parks, Paul A Raschky, and Michael J Tierney, "African leaders and the geography of China's foreign assistance," *Journal of Development Economics*, 2019, *140*, 44–71.

Elvidge, Christopher D, Kimberly Baugh, Mikhail Zhizhin, Feng Chi Hsu, and Tilottama Ghosh, "VIIRS night-time lights," *International journal of remote sensing*, 2017, *38* (21), 5860–5879.

Eurostat, "GISCO: Geographic Information System of the COmmission," 2021. Accessed August 9, 2021, https://ec.europa.eu/eurostat/web/gisco.

Farr, Tom G, Paul A Rosen, Edward Caro, Robert Crippen, Riley Duren, Scott Hensley, Michael Kobrick, Mimi Paller, Ernesto Rodriguez, Ladislav Roth et al., "The shuttle radar topography mission," *Reviews of geophysics*, 2007, *45* (2).

GADM, "Database of Global Administrative Areas," 2018. Accessed August 9, 2021, https://gadm.org/data.html.

Galiani, Sebastian, Paul Gertler, and Ernesto Schargrodsky, "Water for life: The impact of the privatization of water services on child mortality," *Journal of political economy*, 2005, *113* (1), 83–120.

Gething, Peter W, Thomas P Van Boeckel, David L Smith, Carlos A Guerra, Anand P Patil, Robert W Snow, and Simon I Hay, "Modelling the global constraints of temperature on transmission of Plasmodium falciparum and P. vivax," *Parasites & vectors*, 2011, *4* (1), 1–11.

Gilmore, Elisabeth, Nils Petter Gleditsch, Päivi Lujala, and Jan Ketil Rod, "Conflict diamonds: A new dataset," *Conflict Management and Peace Science*, 2005, *22* (3), 257–272.

Glorot, Xavier and Yoshua Bengio, "Understanding the difficulty of training deep feedforward neural networks," in "Proceedings of the thirteenth international conference on artificial intelligence and statistics" JMLR Workshop and Conference Proceedings 2010, pp. 249–256.

Golden, Miriam and Brian Min, "Distributive politics around the world," *Annual Review of Political Science*, 2013, *16* (1), 73–99.

Gorelick, Noel, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, 2017.

Griffiths, Ieuan, "The scramble for Africa: Inherited political boundaries," *The Geographical Journal*, 1986, *152* (2), 204–216.

Guo, Chuan, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger, "On calibration of modern neural networks," in "International Conference on Machine Learning" PMLR 2017, pp. 1321–1330.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in "Proceedings of the IEEE conference on computer vision and pattern recognition" 2016, pp. 770–778.

Henderson, J Vernon, Adam Storeygard, and David N Weil, "Measuring economic growth from outer space," *American economic review*, 2012, *102* (2), 994–1028.

Irwin, Bridget R, Klesta Hoxha, and Karen A Grépin, "Conceptualising the effect of access to electricity on health in low-and middle-income countries: A systematic review," *Global public health*, 2020, *15* (3), 452–473.

Jean, Neal, Marshall Burke, Michael Xie, W Matthew Davis, David B Lobell, and Stefano Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, 2016, *353* (6301), 790–794.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "The worldwide governance indicators: methodology and analytical issues1," *Hague journal on the rule of law*, 2011, *3* (2), 220–246.

Kramon, Eric and Daniel N Posner, "Who benefits from distributive politics? How the outcome one studies affects the answer one gets," *Perspectives on Politics*, 2013, *11* (2), 461–474.

Lujala, Päivi, Jan Ketil Rod, and Nadja Thieme, "Fighting over oil: Introducing a new dataset," *Conflict Management and Peace Science*, 2007, *24* (3), 239–256.

Marshall, Monty G, Ted Robert Gurr, and Keith Jaggers, "Polity IV project: Political regime characteristics and transitions, 1800–2018," *Center for Systemic Peace*, 2019.

McCauley, John F and Daniel N Posner, "African borders as sources of natural experiments promise and pitfalls," *Political Science Research and Methods*, 2015, *3* (2), 409–418.

Meijer, Johan R, Mark AJ Huijbregts, Kees CGJ Schotten, and Aafke M Schipper, "Global patterns of current and future road infrastructure," *Environmental Research Letters*, 2018, *13* (6), 064006.

Michalopoulos, Stelios and Elias Papaioannou, "Pre-colonial ethnic institutions and contemporary African development," *Econometrica*, 2013, *81* (1), 113–152.

_ and _, "National institutions and subnational development in Africa," *The Quarterly journal of economics*, 2014, *129* (1), 151–213.

Moscona, Jacob, Nathan Nunn, and James A Robinson, "Segmentary Lineage Organization and Conflict in Sub-Saharan Africa," *Econometrica*, 2020, *88* (5), 1999–2036.

Murdock, George Peter, "Africa its peoples and their culture history," 1959.

Nunn, Nathan, "The long-term effects of Africa's slave trades," *The Quarterly Journal of Economics*, 2008, *123* (1), 139–176.

Oshri, Barak, Annie Hu, Peter Adelson, Xiao Chen, Pascaline Dupas, Jeremy Weinstein, Marshall Burke, David Lobell, and Stefano Ermon, "Infrastructure quality assessment in africa using satellite imagery and deep learning," in "Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining" 2018, pp. 616–625.

Pande, Rohini and Christopher Udry, "Institutions and development: A view from below," *Yale University Economic Growth Center Discussion Paper*, 2005, (928).

Pebesma, Edzer, "Simple Features for R: Standardized Support for Spatial Vector Data," *The R Journal*, 2018, *10* (1), 439–446.

Ramankutty, Navin, Jonathan A Foley, John Norman, and Kevin McSweeney, "The global distribution of cultivable lands: current patterns and sensitivity to possible climate change," *Global Ecology and biogeography*, 2002, *11* (5), 377–392.

Ratledge, Nathan, Gabe Cadamuro, Brandon de la Cuesta, Matthieu Stigler, and Marshall Burke, "Using machine learning to assess the livelihood impact of electricity access," *Nature*, 2022, *611* (7936), 491–495.

Saavedra, Santiago and Mauricio Romero, "Local incentives and national tax evasion: The response of illegal mining to a tax reform in Colombia," *European Economic Review*, 2021, *138*, 103843.

Tiecke, Tobias G, Xianming Liu, Amy Zhang, Andreas Gros, Nan Li, Gregory Yetman, Talip Kilic, Siobhan Murray, Brian Blankespoor, Espen B Prydz et al., "Mapping the world population one building at a time," *arXiv preprint arXiv:1712.05839*, 2017.

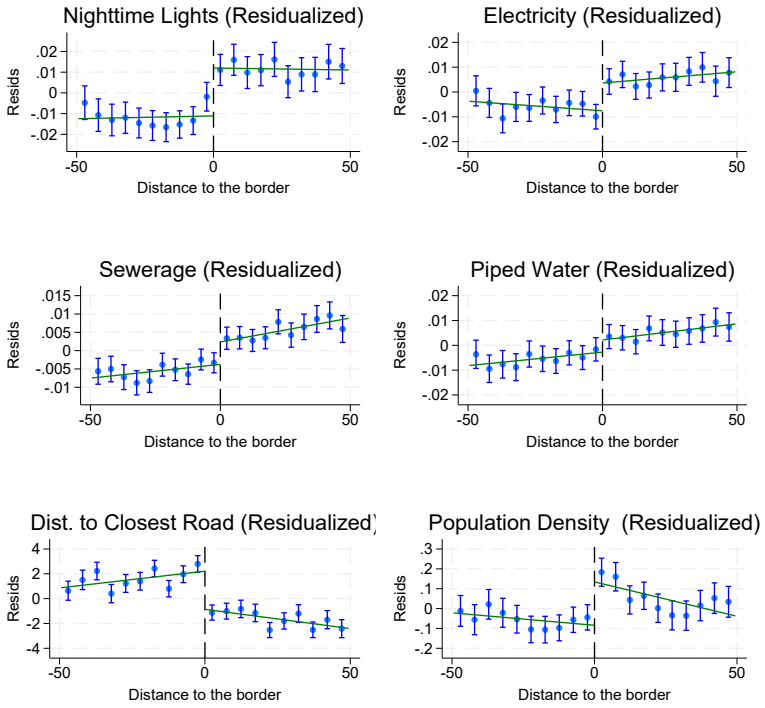
Xie, Michael, Neal Jean, Marshall Burke, David Lobell, and Stefano Ermon, "Transfer learning from deep features for remote sensing and poverty mapping," 2016, *30* (1).

Yeh, Christopher, Anthony Perez, Anne Driscoll, George Azzari, Zhongyi Tang, David Lobell, Stefano Ermon, and Marshall Burke, "Using publicly available satellite imagery and deep learning to understand economic well-being in Africa," *Nature communications*, 2020, *11* (1), 1–11.

Zizka, Alexander, Daniele Silvestro, Tobias Andermann, Josue Azevedo, Camila Duarte Ritter, and Daniel Edler, "Package 'CoordinateCleaner'," 2019.

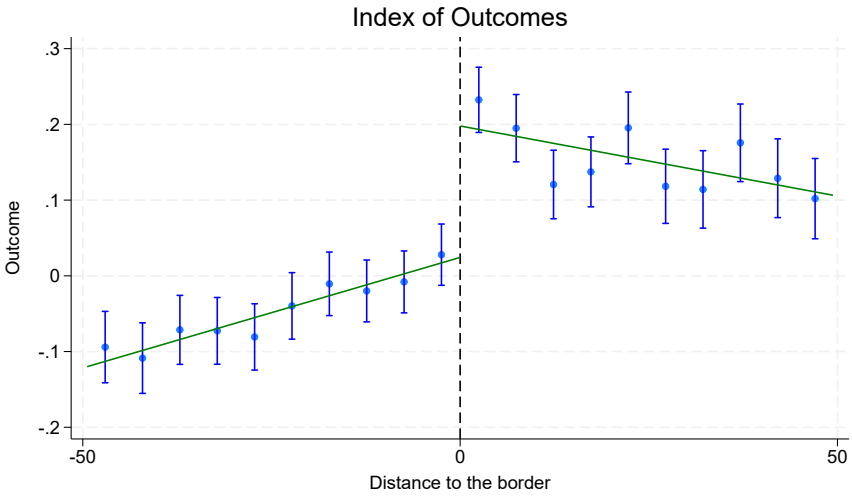
A Appendix Figures

Figure A.1: Discontinuity around national borders, by outcome (residualized)



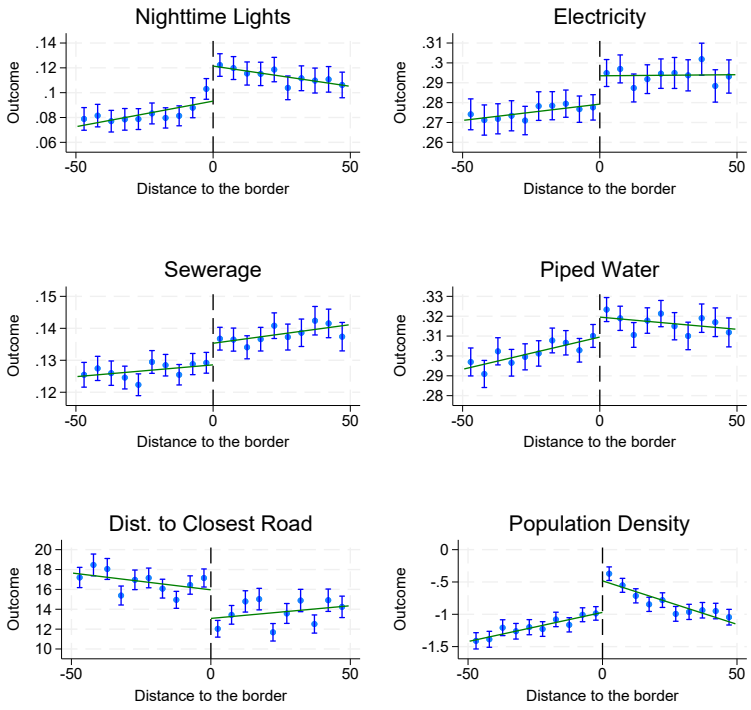
Note: The country with the better institutions is on the right (positive distance to the border).

Figure A.2: Discontinuity around national borders (not residualized)



Note: The country with the better institutions is on the right (positive distance to the border).

Figure A.3: Discontinuity around national borders, by outcome (not residualized)



Note: The country with the better institutions is on the right (positive distance to the border).

Figure A.4: National Institutions: Robustness to polynomial degree choice

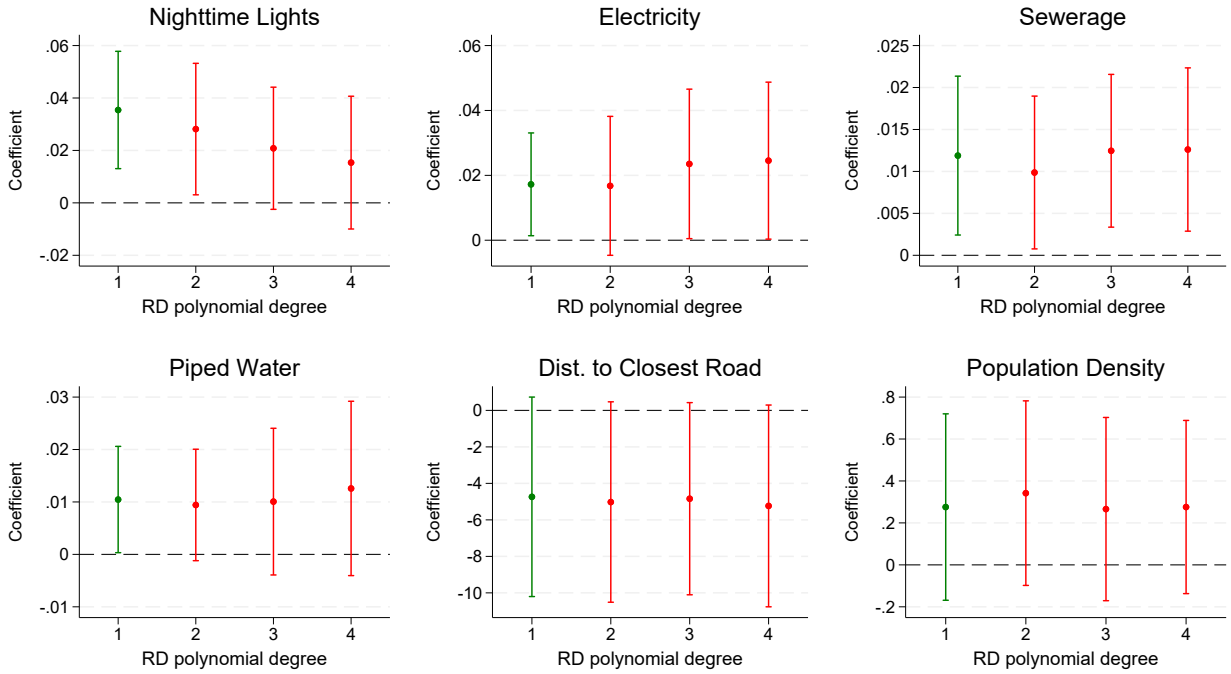


Figure A.5: Gaps in institutions across national borders

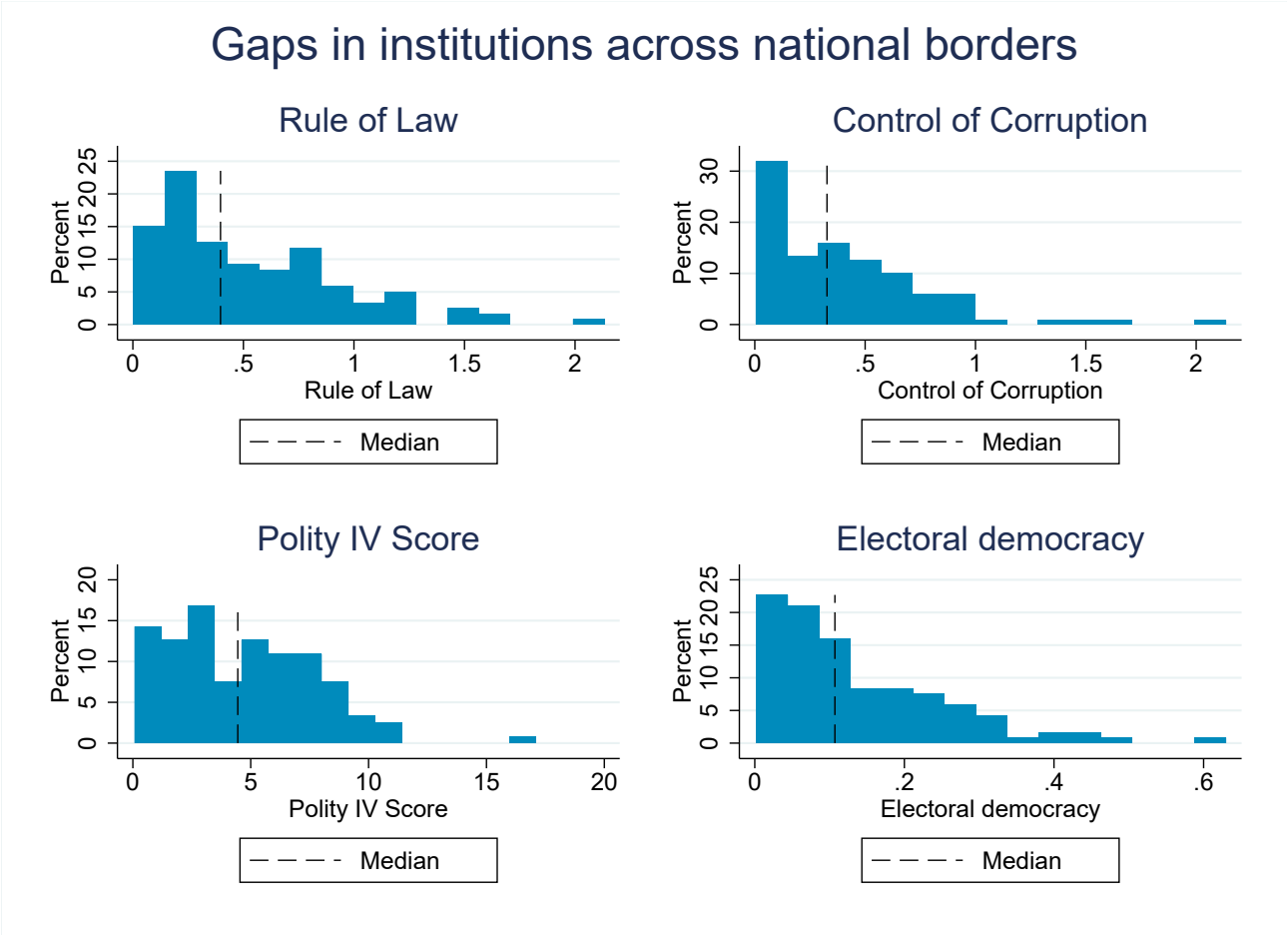


Figure A.6: National Institutions: Robustness to bandwidth choice

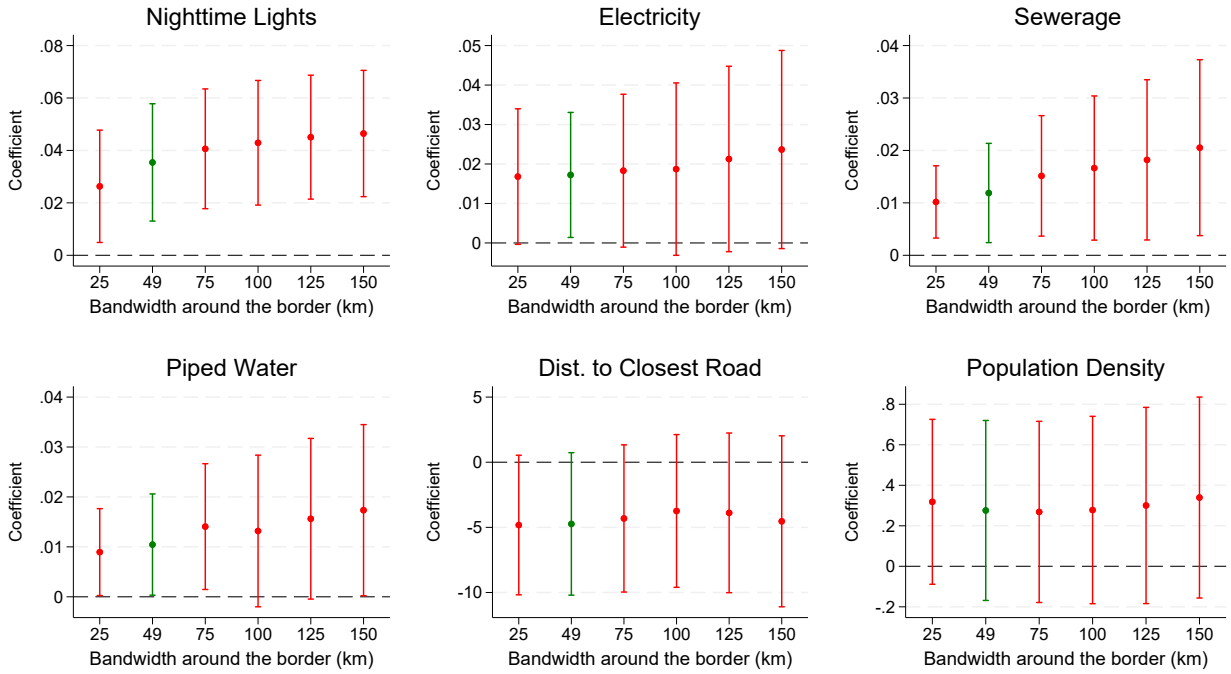
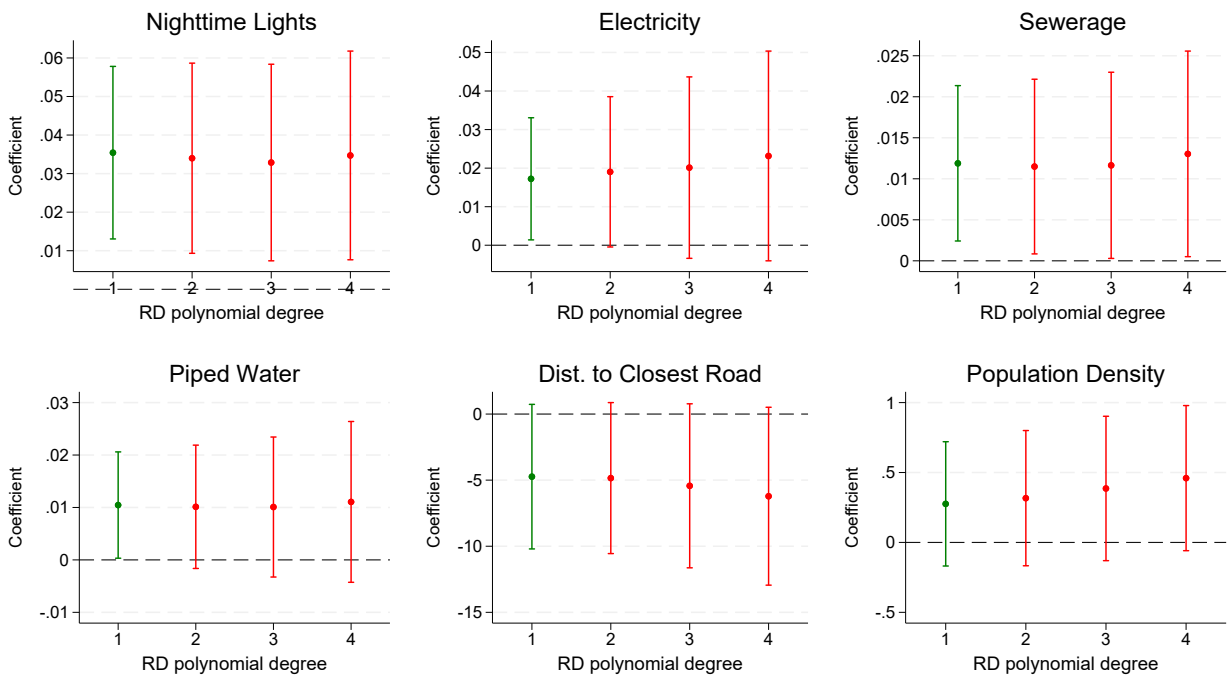


Figure A.7: National Institutions: Robustness to polynomial degree with optimal bandwidth choice



B Appendix Tables

Table B.1: Model Performance for Infrastructure Outcomes: Urban vs. Rural

	Mean	Accuracy	Recall	Specificity	AUCROC
Electricity					
Urban	0.934	0.845	0.866	0.545	0.706
Rural	0.328	0.824	0.753	0.859	0.806
Piped water					
Urban	0.863	0.818	0.865	0.522	0.693
Rural	0.324	0.745	0.667	0.782	0.725
Sewerage					
Urban	0.642	0.713	0.721	0.700	0.710
Rural	0.069	0.922	0.488	0.955	0.721

Note: Urban definition is from Afrobarometer survey. Desert EAs are counted as rural. "Accuracy" is the share of predictions that are correct. XX add definitions for recall, specificity and AUCROC XXX

Table B.2: Datasets for the validation exercises

Survey	Year	Landsat Period	Nightlight Period	Exporting Method
Training Dataset				
Afrobarometer R6 (multi-country)	2014-2015	2014-2015	2015-2016	centroid
Validation datasets: Same Time Period, Different Levels				
Household Access				
Neno & Mwanza, Malawi (rural)	2015	2014-2015	2015-2016	patch level
Greater Abidjan, Côte d'Ivoire (urban)	2014	2014-2015	2015-2016	patch level
Greater Addis Ababa, Ethiopia (urban)	2016	2014-2015	2015-2016	patch level
Institutional Access				
Ghana schools (country-wide)	2015	2014-2015	2015-2016	centroid
Validation datasets: Different Time Periods				
Afrobarometer R5 (multi-country)	2011-2013	2011-2013	2012-2014	centroid
Greater Addis Ababa, Ethiopia (urban)	2018	2018	2018	patch level

Table B.3: Validation Results

Dataset	**Benchmark-NL**			**Model - R6**		
	Electricity	Piped Water	Sewerage	Electricity	Piped Water	Sewerage
Afrobarometer R5						
All	0.59	0.56	0.28	0.74	0.66	0.79
Urban	0.92	0.85	0.61	0.84	0.76	0.66
Rural	0.40	0.40	0.09	0.69	0.59	0.86
Border	0.42	0.48	0.15	0.69	0.60	0.83
Ghana						
All	0.89	0.50		0.88	0.58	
Urban	0.97	0.58		0.94	0.62	
Rural	0.83	0.44		0.82	0.55	
Border	0.89	0.28		0.85	0.34	
Other surveys						
Addis - W1	0.84	0.65	0.22	0.71	0.52	0.71
Addis - W2	0.75	0.47		0.67	0.58	
Abidjan	0.80	0.54	0.05	0.75	0.58	0.76
Malawi (Rural)	0.56	0.56	0.53	0.95	0.95	1.00

Note: On the first 3 columns, we display the accuracy on different validation datasets of a "naive" model based in nighttime lights: we predict that there is access to infrastructure if nighttime lights are positive. The next 3 columns show the accuracy of our convolutional neural network model trained with Afrobarometer R6 groundtruth and satellite imagery.

Table B.4: Extra validation Results

Dataset	Benchmark-NL			Model - R6		
	Electricity	Piped Water	Sewerage	Electricity	Piped Water	Sewerage
Ghana (Threshold: 0)						
All	0.89	0.50		0.88	0.58	
Urban	0.97	0.58		0.94	0.62	
Rural	0.83	0.44		0.82	0.55	
Border	0.89	0.28		0.85	0.34	
Other surveys (Threshold: 0)						
Addis - W1	0.84	0.65	0.22	0.71	0.52	0.71
Addis - W2	0.75	0.47		0.67	0.58	
Abidjan	0.80	0.54	0.05	0.75	0.58	0.76
Malawi (Rural)	0.56	0.56	0.53	0.95	0.95	1.00
Ghana (Threshold: 0.2)						
All	0.89	0.50		0.88	0.58	
Urban	0.97	0.58		0.94	0.62	
Rural	0.83	0.44		0.82	0.55	
Border	0.89	0.28		0.85	0.34	
Other surveys (Threshold: 0.2)						
Addis - W1	0.84	0.65	0.22	0.71	0.52	0.71
Addis - W2	0.75	0.47		0.67	0.58	
Abidjan	0.80	0.54	0.05	0.75	0.58	0.76
Malawi (Rural)	0.56	0.56	0.53	0.95	0.95	1.00

Note: On the first 3 columns, we display the accuracy on different validation datasets of a "naive" model based in nighttime lights: we predict that there is access to infrastructure if nighttime lights are positive. The next 3 columns show the accuracy of our convolutional neural network model trained with Afrobarometer R6 groundtruth and satellite imagery.

Table B.5: Distribution of Infrastructure Measures Across Countries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome	**Index**	**Nightlights**	**Electricity**	**Sewerage**	**Piped water**	**Dist. Road**	**Pop. Density**
Mean	0.35	0.20	0.37	0.16	0.38	56.15	0.02
Std. Dev.	2.45	0.22	0.19	0.06	0.15	284.30	2.94
Minimum	-15.64	0.00	0.09	0.08	0.12	0.05	-4.61
Percentile 25	-0.09	0.05	0.24	0.12	0.29	0.91	-2.58
Median	0.62	0.10	0.36	0.14	0.37	4.00	0.26
Percentile 75	1.26	0.29	0.46	0.17	0.41	15.29	2.24
Maximum	3.93	0.96	0.89	0.36	0.84	2101.24	5.45
Interquartile Range	1.35	0.24	0.22	0.05	0.13	14.37	4.82
Coefficient/IQR	0.14	0.15	0.08	0.25	0.08	-0.33	0.06

Notes: We report country level summary statistics. Coefficient/IQR ratios are defined relative to coefficients in Panel A of Table 3.

Table B.6: Optimal bandwidth for different polynomial degrees

Polynomial degree	**Optimal bandwidth (km)**
1	49.5
2	82.9
3	128.7
4	179.5

Table B.7: Accounting for multiple hypothesis testing

Outcome	**Night lights**	**Electricity**	**Sewerage**	**Piped water**	**Dist. Road**	**Pop. Density**
Panel A. Index of institutions:						
Coefficient	0.035	0.017	0.012	0.010	-4.736	0.276
Normal p-value	0.000	0.000	0.000	0.000	0.000	0.000
Clustered p-value	0.003	0.034	0.015	0.044	0.088	0.218
Clustered-FDR p-value	0.016	0.048	0.039	0.048	0.070	0.079
Panel B. Rule of Law:						
Coefficient	0.036	0.007	0.008	0.010	6.192	0.162
Normal p-value	0.000	0.005	0.000	0.000	0.000	0.000
Clustered p-value	0.000	0.313	0.038	0.035	0.016	0.446
Clustered-FDR p-value	0.002	0.144	0.051	0.051	0.042	0.175
Panel C. Control of Corruption:						
Coefficient	0.030	0.006	0.007	0.008	-0.947	0.049
Normal p-value	0.000	0.013	0.000	0.001	0.003	0.137
Clustered p-value	0.002	0.316	0.040	0.070	0.686	0.806
Clustered-FDR p-value	0.015	0.311	0.110	0.132	0.676	0.676
Panel D. Polity IV Score:						
Coefficient	0.033	0.017	0.006	0.007	-6.155	-0.093
Normal p-value	0.000	0.000	0.000	0.002	0.000	0.004
Clustered p-value	0.008	0.016	0.083	0.125	0.015	0.727
Clustered-FDR p-value	0.034	0.034	0.067	0.081	0.034	0.176
Panel E. Electoral Democracy:						
Coefficient	0.042	0.017	0.006	0.010	-4.445	0.580
Normal p-value	0.000	0.000	0.000	0.000	0.000	0.000
Clustered p-value	0.001	0.031	0.118	0.061	0.102	0.002
Clustered-FDR p-value	0.006	0.043	0.066	0.066	0.066	0.006

Table B.8: Population density results

Outcome	PD 2000	PD 2015 (GRIP)	PD 2015 (Meta)
Panel A. Main specification:			
Index of institutions (PCA)	-0.2084**	-0.1071	0.2876
	(0.089)	(0.117)	(0.212)
	{0.02}	{0.37}	{0.18}
Observations	74,843	77,712	77,743
Panel B: breakdown by measure of institutional quality			
Rule of Law	-0.2280**	-0.1416	0.1602
	(0.089)	(0.097)	(0.215)
	{0.01}	{0.15}	{0.46}
Control of Corruption	-0.1460*	-0.1840**	0.1341
	(0.078)	(0.090)	(0.205)
	{0.07}	{0.05}	{0.52}
Polity IV Score	-0.3133***	-0.1801	-0.0784
	(0.099)	(0.110)	(0.262)
	{0.00}	{0.11}	{0.77}
Electoral democracy	-0.3430***	-0.1171	0.5901***
	(0.077)	(0.110)	(0.168)
	{0.00}	{0.29}	{0.00}
Observations	74,843	77,712	77,743
Panel C. No controls:			
Index of institutions (PCA)	-0.1708	-0.1273	0.5565**
	(0.119)	(0.114)	(0.222)
	{0.16}	{0.27}	{0.02}
Observations	74,843	77,712	77,743
Panel D. Large discontinuities:			
Index of institutions (PCA)	-0.5266***	-0.2958**	0.1627
	(0.116)	(0.130)	(0.164)
	{0.00}	{0.03}	{0.33}
Observations	42,000	42,133	42,150
Panel E. Ethnicities with the same colonizer:			
Index of institutions (PCA)	-0.5764***	-0.6354***	-0.5720**
	(0.145)	(0.155)	(0.211)
	{0.00}	{0.00}	{0.01}
Observations	26,645	28,102	28,125
Panel F. Neither patch within 300 km of the capital:			
Index of institutions (PCA)	-0.1112	-0.1016	0.2600
	(0.115)	(0.126)	(0.247)
	{0.34}	{0.43}	{0.30}
Observations	40,978	42,634	42,651

Notes: We include ethnicity level fixed effects, standard errors are clustered at the country level and are reported in parentheses, p-values are reported in curly braces.

Table B.9: Robustness for Political Ties (Above Median Duration)

Outcome	Index	Nightlights	Electricity	Sewerage	Piped water	Dist. Road
Political tie above median	1.4373***	1.0853***	0.1181***	0.1559***	0.1290***	-0.4568*
	(0.289)	(0.314)	(0.028)	(0.038)	(0.030)	(0.244)
	{0.00}	{0.00}	{0.00}	{0.00}	{0.00}	{0.07}
Political tie above median × Index of institutions	-0.3289**	-0.1468	-0.0325**	-0.0330	-0.0236*	0.0922
	(0.160)	(0.146)	(0.015)	(0.020)	(0.014)	(0.133)
	{0.05}	{0.32}	{0.04}	{0.10}	{0.09}	{0.49}
Political tie below median	0.7276***	0.2538***	0.0817***	0.0508**	0.0721***	-0.5363
	(0.175)	(0.067)	(0.021)	(0.020)	(0.019)	(0.325)
	{0.00}	{0.00}	{0.00}	{0.02}	{0.00}	{0.11}
Political tie below median × Index of institutions	-0.0096	0.0076	-0.0012	0.0038	0.0016	0.0576
	(0.104)	(0.077)	(0.012)	(0.013)	(0.010)	(0.227)
	{0.93}	{0.92}	{0.92}	{0.77}	{0.88}	{0.80}
Outcome mean	1.88	0.48	0.46	0.23	0.44	1.35
Observations	4,633	4,633	4,633	4,633	4,633	4,633
$\boldsymbol{R^2}$	0.59	0.61	0.53	0.65	0.49	0.27

Notes: The median duration of leaders in power in our data is 6 years. Our African political leaders' birthregion dataset covers 283 leaders from 46 countries. It includes birth regions and ethnicity information on political leader since their countries' independence. Political ties (above median duration) are equal to 1 when the place is associated with a leader who was in power for over 6 years.

C Appendix: Technical Notes on CNNs and Calibration

C.1 Machine Learning Model

We build on the approach developed in Oshri et al. (2018). The first task consists in solving a multi-label binary classification problem: Using as a input a satellite image X, we want to predict whether a given enumeration area (our observation level) has or not access to infrastructure defined by a set of binary label $Y_1, Y_2, \ldots, Y_k \in \{0, 1\}$. For each infrastructure access label Y_i we train a separate binary classifier $h_i(X) = \{\hat{Y}_i, \hat{P}_i\}$, where $\hat{Y}_i$ is a class prediction and $\hat{P}_i$ is the confidence associated to that prediction.

To tackle our binary image classification problem, we use deep learning techniques—algorithms that combine a large amount of linear and non-linear functions in a hierarchical way, making the algorithms very flexible, and thus allowing them to learn very complicated functional forms. Convolutional Neural Networks (CNNs) are a class of deep learning algorithms used for image classification tasks, with their main advantages being that they preserve the spatial structure of images, and can learn patterns contained in images with fewer parameters than a regular neural network. The hierarchical structure of CNNs allows the algorithms to extract basic or low-level features from data, and then to aggregate these basic features to learn more complex features. As the algorithms identify features without supervision, researchers no longer need to program complex routines to identify edges or shapes.

We train a Residual Network (ResNet), a state-of-art CNN that has been shown to have good performance for binary classification problems (He et al., 2016). We modify the structure of the initial 18-layer Residual Network in 2 ways: we start by increasing the number of input channels in the first convolutional layer, given that the original model was trained to classify 3 band images on ImageNet and we are using satellite imagery with six bands for the Landsat data and 1 band for the VIIRS Nighttime data. We also modify the final fully connected softmax layer of the model, that was originally meant to classify an image into 1,000 categories, to work with a binary classification problem and predict the probability of an enumeration area having access to an infrastructure variable. We initialize the weights for the RGB bands of the Landstat images using the the ResNet pre-trained weights, and we use Xavier initialization to initialize the weights of remaining Landsat and VIIRS Nighttime bands, given that drawing initial values from the uniform distribution described in Glorot and Bengio (2010) has been shown to lead to faster convergence. We follow the same strategy regarding data processing, regularization and optimization as Oshri et al. (2018).

C.2 Calibration

Calibration consists in modifying the "confidence" ($\hat{p}_i$) produced by the model, in order for the distribution of generated confidences to resemble the empirical distribution in the training dataset, without affecting the accuracy of the model. Current deep learning models tend to generate confidence levels that are very close to 0 or 1 for each observation, so they can get a small loss function. As economists or political scientists, however, we care about the *likelihood* that a given area has access to specific amenities—rather than the binary prediction, to the extent that it is correlated with the level of development. In other words, if the model predicts

that area A has electricity with probability 55%, area B has electricity with probability 65% and area C with 85%, while all three areas would have a binary prediction of "1", we want to be able to interpret the A-C gap of 30 percentage points as having some cardinal interpretation: area C is"more developed" than area A, and the gap in development between C and A is greater than that between B and A. Calibration helps ensure that the confidence (the predicted likelihood) can be interpreted in that way.

We use a post-estimation calibration technique called "temperature scaling" (Guo et al., 2017), that works in the following way: After estimating the parameters in the Residual Network $\hat{h}_i$, we compute scores $\hat{s}_i$ with the data in our validation set. The scores are the inputs to the final softmax layer of our network that generates the confidence or predicted probability $\hat{p}_i$. We then run a logistic regression of the true classes in our validation set, $\hat{y}_i$, against our score, of the form $\hat{y}_i = \sigma(\frac{\hat{s}_i}{T} + \epsilon)$, where σ is the logistic function. The newly estimated temperature parameter is the logistic function. The newly estimated temperature parameter $\hat{T}$ is then used to adjust the confidence, so now we get the new calibrated prediction $\tilde{p}_i \equiv \sigma(\frac{\hat{s}_i}{\hat{T}})$

D Appendix: Dataset building

D.1 Grid of patches, national borders, valid patches

To build our dataset, we start by generating a grid of 6.72 by 6.72 km patches over all the African continent. We built a map of the continent of Africa by aggregating country shapefiles obtained from GADM (GADM, 2018) and GISCO (Eurostat, 2021), and then we divided it in 747,165 patches using the SF package for R (Pebesma, 2018).

We add to our grid the shapefiles containing the borders of the ethnical ethnicities of Africa described in Murdock (1959), and we also add individual shapefiles for every African country from GADM. We intersect both maps to identify ethnicities that are partitioned by national borders.

Not every patch of land is useful for our analysis. First, we exclude patches that are in the border of 2 ethnicities or 2 countries. We also use a water cover raster from the MOD44W MODIS dataset (Carroll et al., 2017) and a forest cover raster from the Dynamic Land Cover map from (Buchhorn et al., 2020). Patches marked as part of a forest or a water body are not considered for our analysis, since they will mechanically have no access to infrastructure, and this is caused by geographical factor rather than by national or pre-colonial institutions.

When studying ethnicities that are partitioned by national borders, it might be the case that more than one country can be present in a partitioned ethnicity (for instance, the territory of the Malinke ethnicity falls within 6 countries), and it there might also be discontinuities in the distribution of the distance to the border within certain ethnicities caused by ethnicities and countries that intersect in 2 separate areas, or parts of the territory of a ethnicity being mostly covered by a forest or a water body, so in order to overcome this, we define a set of criteria to assign patches in each ethnicity to the main border within the ethnicity and compute the distance to the border to that border. To do this, we use the following iterative procedure:

1. We identify all the feasible borders within a ethnicity.
2. We define the main border as the border separating 2 adjacent countries that combined have the biggest share of the area of the ethnicity.
3. For each patch we compute the distance to the main border.
4. We check for discontinuities in the distribution of distance to the border. If there is a gap of more than 10 km, all the patches outside the gap are no longer considered as belonging to our previous main border, but they can be considered as a part of a different border.
5. If one of the sides of the ethnicity has less than 10 patches, or if after removing patches due to a discontinuity we have another border covering a bigger area, we replace our main border with the new biggest border and go back to step 3, and we repeat this until there are no changes and we have a stable main border.

D.2 Outcomes and controls

We intersected our grid with several rasters of data, and we also added a lot of datasets to include the following variables:

- **Malaria temperature suitability**: We obtained data for a temperature suitability for P. Falciparum and P. Vivax transmission for 2010 from Gething et al. (2011). Their data is at a resolution of 1 km by 1 km, so for each of our patches we take the average value of all the pixels that fall inside them using Google Earth Engine (Gorelick et al., 2017).
- **Land suitability**: We obtained a measure of land suitability for agriculture from The Atlas of the Biosphere (Ramankutty et al., 2002). This dataset has information about the fraction of each cell that is suitable to be used for agriculture. The original raster file is at a resolution of 0.5 degrees (approximately 55 kilometers around the Equator), so we intersect our patches with this raster, and take the average suitability for the cases where our patches lie within more than one cells of the original raster.
- **Elevation**: We use elevation data from the Shuttle Radar Topography Mission from NASA (Farr et al., 2007) with a 30 meters resolution. For each of our patches of land we take the average value of the raster with elevation using Google Earth Engine.
- **Distance to the equator**: To control for the distance to the equator, we take the absolute value of the latitude of each patch.
- **Distance to the capital**: We compute our distance to the capital for each country by retrieving the geographical coordinates of their capital using the R package CoordinateCleaner (Zizka et al., 2019), and then we compute the distance to the capital from each patch using the SF package (Pebesma, 2018).
- **Distance to the coast**: To define our distance to the coast, we combine the GADM maps of all the African countries to obtain an outline of the African continent, and then we compute the distance from each patch to the coast using the SF package.
- **Distance to petroleum sources**: We identify the location of petroleum sources with the PETRODATA v1.2 dataset from the Peace Research Institute Oslo (Lujala et al., 2007). For each patch, we again use the SF package to find the closest petroleum sources inside and outside of the country the patch belongs to, and then we compute the distance to these 2 petroleum sources.
- **Distance to diamond sources**: We identify the location of diamond sources by using the DIADATA v1a dataset from the Peace Research Institute Oslo (Gilmore et al., 2005). For each patch, we use the SF package to identify the closest diamond sources inside and outside of the country the patch belongs to, and then we compute the distance to these 2 diamond sources.
- **Distance to the closest road**: We use the GRIP global roads database to compute the distance to the closest road for each patch of land in our grid (Meijer et al., 2018). This dataset contains maps for all the type of roads (highways, primary, secondary, tertiary and local roads) for all the continents, so we intersect this with country maps for Africa, and for each patch we compute the distance to the closest road within their respective country using the SF package in R.
- **Nightlights**: We use VIIRS Nighttime Light data (Elvidge et al., 2017). Specifically, we use annual composites for the year 2015 of the version 1, where they preprocess imagery to remove stray lights and outliers, average the annual data using only cloud-free luminosity

measurements, and sets the background lights to 0. These images have a resolution of 15 arc seconds (around 500m at the Equator), so we use Google Earth Engine to compute the average luminosity for each of our patches of land.

- **Population density**: We have 2 data sources for population density for the year 2015: We use the Gridded Population of the World v4 dataset (Center for International Earth Science Information Network, CIESIN, 2016). This data has a resolution of 30 seconds (approximately 1 km around the equator), so here we again take averages of the values of the pixels that are inside each of the patches in our grid using Google Earth Engine. We also use the Data For Good at Meta's High Resolution Population Density Maps (Tiecke et al., 2017), a product based on the Gridded Population of the World v4 dataset that uses computer vision algorithms to re-distribute population density within territories according to the volume of the buildings in said territory. These predictions have a 30 meter resolution, so for each patch of land in our grid we take the average value of the pixels in the population density raster using Google Earth Engine.

E ML Model Evaluation

F Colonizers and rulers at independence

The identification strategy in our first application relies on working with ethnicities that were partitioned by national borders, so that we can compare territories that have common ethnic origins and should not be too different apart from the different national institutions they are exposed to. Despite this, we could also think that European colonizers shaped the early institutions of modern African countries, so ethnicities that were separated into countries with different colonizers could have underlying differences that can be confounded for differences in national institutions.

To overcome this issue we produced a list of colonizers for each African country, and the we use this to re-estimate our results using a sample of ethnicities that were partitioned into countries with the same colonizers, or adding colonizer fixed effects. After World War I the allies re-distributed the control of some African colonies, so as an extra robustness check we also compiled a list of rulers at independence for most African countries, because rulers in the late colonial days might have been more important than early colonizers. In table D.1 we present a list of African countries, their colonizers and their rulers at the time of their independence. For some countries we couldn't establish a single colonizer or ruler at independence: for instance, Somalia was controlled by Great Britain and Italy, so we cannot assign it to a single colonizer or ruler at independence. We code these colonizer and ruler into 3 categories:

Table D.1: Colonizers and rulers at independence in Africa

Country	Colonizer	Ruler at independence	Country	Colonizer	Ruler at independence
Algeria	France	France	Madagascar	France	France
Angola	Portugal	Portugal	Malawi	Great Britain	Great Britain
Benin	France	France	Mali	France	France
Botswana	Great Britain	Great Britain	Mauritania	France	France
Burkina Faso	France	France	Mauritius	France	Great Britain
Burundi	Germany	Belgium	Mozambique	Portugal	Portugal
Côte d'Ivoire	France	France	Namibia	Germany	Germany
Cameroon	Germany	-	Niger	France	France
Cape Verde	Portugal	Portugal	Nigeria	-	Great Britain
Central African Republic	France	France	Republic of Congo	France	France
Chad	France	France	Rwanda	Germany	Belgium
Comoros	France	France	São Tomé and Príncipe	Portugal	Portugal
Democratic Republic of the Congo	Belgium	Belgium	Saint Helena	Great Britain	Great Britain
Djibouti	France	France	Senegal	France	France
Egypt	Great Britain	Great Britain	Seychelles	Great Britain	Great Britain
Equatorial Guinea	Spain	Spain	Sierra Leone	Great Britain	Great Britain
Eritrea	Italy	Great Britain	South Africa	Great Britain	Great Britain
Ethiopia	Italy	Italy	South Sudan	Great Britain	Great Britain
Gabon	France	France	Sudan	Great Britain	Great Britain
Gambia	-	Great Britain	Swaziland	Great Britain	Great Britain
Ghana	-	Great Britain	Tanzania	Germany	Great Britain
Guinea	France	France	Togo	Germany	-
Guinea-Bissau	Portugal	Portugal	Tunisia	France	France
Kenya	-	Great Britain	Uganda	Great Britain	Great Britain
Lesotho	Great Britain	Great Britain	Zambia	Great Britain	Great Britain
Liberia	United States	United States	Zimbabwe	Great Britain	Great Britain
Libya	Italy	-			

With these definitions we replicate our main results restricting our sample to ethnicities with the same colonizers or the same rule at independence, or using our full sample with colonizer or ruler at independence fixed effects. We only include fixed effects for France and Great Britain, since all other colonizers and rulers combined controlled less than 25% of the territory in Africa. In Table D.2 we present the results of these experiments: our results remain fairly similar to the main results presented in Panel A of Table 3, with the coefficients for Nighttime lights, Electricity, Sewerage, Piped ater and our PCA index being significant in most specifications, and presenting coefficients with similar magnitudes to the ones we observed in our main results.

Table D.2: Alternative RD specifications

Outcome	Index	Nightlights	Electricity	Sewerage	Piped water	Dist. Road	Pop. Density
Panel A. Ethnicities with the same colonizer:							
Index of institutions (PCA)	0.1312*	0.0467**	0.0317***	0.0182**	0.0077	-0.5426	-0.6146***
	(0.065)	(0.020)	(0.010)	(0.008)	(0.009)	(4.022)	(0.212)
	{0.05}	{0.02}	{0.00}	{0.03}	{0.40}	{0.89}	{0.01}
Observations	29,667	29,673	29,667	29,667	29,667	29,673	29,673
Panel B. Colonizer fixed effects:							
Index of institutions (PCA)	0.2085***	0.0396***	0.0184**	0.0125***	0.0119**	-5.3735**	0.2449
	(0.049)	(0.010)	(0.008)	(0.005)	(0.005)	(2.553)	(0.211)
	{0.00}	{0.00}	{0.02}	{0.01}	{0.02}	{0.04}	{0.25}
Observations	82,082	82,088	82,082	82,082	82,082	82,088	82,088
Panel C. Ethnicities with the same ruler at independence:							
Index of institutions (PCA)	0.0939	0.0417**	0.0310***	0.0173**	0.0019	3.0639	-0.5296**
	(0.066)	(0.019)	(0.009)	(0.008)	(0.009)	(3.873)	(0.216)
	{0.17}	{0.04}	{0.00}	{0.03}	{0.84}	{0.44}	{0.02}
Observations	30,753	30,759	30,753	30,753	30,753	30,759	30,759
Panel D. Ruler at independence fixed effects:							
Index of institutions (PCA)	0.2004***	0.0379***	0.0176**	0.0122**	0.0110**	-4.6497*	0.2779
	(0.052)	(0.010)	(0.008)	(0.005)	(0.005)	(2.704)	(0.212)
	{0.00}	{0.00}	{0.03}	{0.01}	{0.03}	{0.09}	{0.20}
Observations	82,082	82,088	82,082	82,082	82,082	82,088	82,088

Notes: We include ethnicity level fixed effects, standard errors are clustered at the country level and are reported in parentheses, p-values are reported in curly braces.