Supplementary Materials

Economic development in pixels: The limitations of nightlights and new spatially disaggregated measures of consumption and poverty

John D. Huber
Columbia University
Laura Mayoral
Institute for Economic Analysis and Barcelona School of Economics

June 2023

This document provides supplementary materials for the analysis presented in "Economic development in pixels..." Section A presents a table that provides a list of all surveys used in constructing the DHS training variable. Section B is a table that describes the variables used in the prediction exercise. Section C describes the process used for tuning the random forest models, and provides a table listing the final parameters selected. Section D provides additional information and estimation results regarding the analysis of the models in Michalopoulos and Papaioannou (2013 and 2014) using the SED data.

A. DHS surveys used to construct the training variable.

Country	Year	Country	Year
Angola	2006	Mozambique	2009
Angola	2011	Mozambique	2011
Angola	2016	Mozambique	2015
Benin	2012	Mozambique	2018
Benin	2017	Namibia	2007
Burkina Faso	2010	Namibia	2013
Burkina Faso	2014	Nigeria	2008
Burkina Faso	2018	Nigeria	2010
Burundi	2010	Nigeria	2013
Burundi	2012	Nigeria	2015
Burundi	2016	Nigeria	2018
Cameroon	2011	Rwanda	2008
Chad	2015	Rwanda	2010
Dem. Rep. of Congo	2007	Rwanda	2015
Dem. Rep. of Congo	2013	Senegal	2008
Ethiopia	2011	Senegal	2011
Ethiopia	2016	Senegal	2013
Gabon	2016	Senegal	2014
Ghana	2008	Senegal	2015
Ghana	2014	Senegal	2016
Ghana	2016	Sierra Leone	2013
Guinea	2012	Sierra Leone	2016
Guinea	2018	Tanzania	2007
Kenya	2009	Tanzania	2010
Kenya	2014	Tanzania	2012
Kenya	2015	Tanzania	2015
Lesotho	2009	Tanzania	2017
Lesotho	2014	Togo	2014
Liberia	2007	Togo	2017
Liberia	2009	Uganda	2006
Liberia	2011	Uganda	2009
Liberia	2013	Uganda	2011
Liberia	2016	Uganda	2014
Madagascar	2011	Uganda	2016
Madagascar	2013	Uganda	2018
Madagascar	2016	Zambia	2007
Malawi	2010	Zambia	2013
Malawi	2012	Zambia	2018
Malawi	2014	Zimbabwe	2010
Malawi	2015	Zimbabwe	2015
Malawi	2017		
Mali	2006		
Mali	2012		
Mali	2015		
Mali	2018		

Table A.1. DHS SURVEYS. This table summarizes the DHS surveys employed in the construction of the training variable.

B. Table describing predictors and their sources

Predictor	Definition	Source
nl_uncal_blur_10	DMSP nightlights (pre-2013) See Elvidege et al (1999) and Hsu et al (2015)	https://eogdata.mines.edu/products/dmsp/#docs
nl_uncal_deblur_10	DMSP deblurred nightlights (pre-2013) See Elvidege et al (1999) and Hsu et al (2015)	https://eogdata.mines.edu/products/dmsp/#docs
nl_viirs_10	VIIRS nightlight measure (post-2012) See Elvidge et al (2013) and Elvidge et al (2017)	https://eogdata.mines.edu/products/vnl/
nls_mean_yeh_c_b	Three-year moving average of NL, DSMP and VIIRS	Following Yeh et al (2020), this variable is created by standardizing nl_uncal_blur_10 and nl_viirs_10 and then taking the cell average of the standardized variable over three-year periods.
d_highway	Distance in meters from cell centroid to nearest highway. See Meijer et al (2018)	https://www.globio.info/download-grip-dataset
d_capital	Distance in meters from cell centroid to national capital	https://hub.arcgis.com/datasets/esri::world-cities/about
d_catholic	Distance in meters from cell centroid to nearest catholic mission	Cagé and Rueda (2020)
d_coast	Distance in meters from cell centroid to nearest coast	https://www.naturalearthdata.com/downloads/10m-physical-vectors/
d_diamonds	Distance in meters from cell centroid to nearest diamond deposit. See Gilmore et al (2005).	https://www.prio.org/data/10

Table continues on next page.

Table describing predictors and their sources, continued

Predictor	Definition	Source
d_harbor	Distance in meters from cell centroid to nearest harbor	http://msi.nga.mil/NGAPortal
d_lakes	Distance in meters from cell centroid to nearest lake	https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database
d_missions	Distance in meters from cell centroid to nearest Christian mission	Nunn (2010)
d_offshoreoil	Distance in meters from cell centroid to nearest off-shore oil and gas deposit. See Lujala et al (2007).	https://www.prio.org/data/11
d_onshoreoil	Distance in meters from cell centroid to nearest on-shore oil and gas deposit. See Lujala et al (2007).	https://www.prio.org/data/11
d_protestant	Distance in meters from cell centroid to nearest protestant mission	Cagé and Rueda (2016)
d_rivers	Distance in meters from cell centroid to nearest river. See Lehner and Döll (2004)	https://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-rivers-lake-centerlines/
d_protected	Distance in meters from cell centroid to nearest protected area	https://www.protectedplanet.net/en/thematic-areas/wdpa?tab=WDPA
remoteness	Predicted score from a PCA that includes d_capital, d_catholic, d_coast, d_diamonds, d_harbor, d_missions, d_offshoreoil, d_onshoreoil, d_protestant, d_highway	
v_temp_avg_10	Average temperature in cell between 1960 and 1990 multiplied by 10. See Harris et al (2020).	https://crudata.uea.ac.uk/cru/data/hrg/
v_malariapf_10	Average prevalence of Malaria Plasmodium falciparum in cell	Weiss et al (2019)

Table continues on next page.

Table describing predictors and their sources, continued

Predictor	Definition	Source
v_rain_10	Average rain in cell in year. See Harris et al (2020).	https://crudata.uea.ac.uk/cru/data/hrg/
v_temp_10	Average temperature in cell in year multiplied by 10. See Harris et al (2020).	https://crudata.uea.ac.uk/cru/data/hrg/
disease	Predicted score from a PCA that includes v_malariapf_10, v_rain_10, v_temp_avg_10	–
v_elevation_10	Average elevation of cell, meters	Berry and Benveniste (2019)
v_calories_10	The average potential yields within each cell attainable given the set of crops	Galor and Özak (2015)
v_rugged_10	Terrain Ruggedness Index quantifying topographic heterogeneity in wildlife habitats providing concealment for preys and lookout posts	Nunn and Puga (2012)
geography	Predicted score from a PCA that includes v_calories_10, v_rugged_10, v_elevation_10	–
v_population_10	Cell population. See Doxsey-Whitfield (2015)	https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-rev11/data-download
v_co2_10	Sum of co2 emissions in each pixel within a square. See EDGARv7.0	https://edgar.jrc.ec.europa.eu/dataset_ghg70
latitude	Latitude of cell centroid	
longitude	Longitude of cell centroid	

Table continues on next page.

Table describing predictors and their sources, continued

Predictor	Definition	Source
eco1	Proportion of cell containing grassland	Olson and Dinerstein (2002)
eco3	Proportion of cell containing grasses and shrubs	Olson and Dinerstein (2002)
eco4	Proportion of cell containing a desert/xeric biome	Olson and Dinerstein (2002)

C. Estimation: Parameter tuning

The main hyperparameters of the models are: the number of individual trees (NTREES), the maximum number of predictors that are included in each tree (NVARS), maximum tree depth (DEPTH), and the minimum proportion of the variance at a node in order for splitting to be performed (VAR). To tune the different models, we consider a grid of values for each of the parameters: 3 values for NTREES, 120, 150 and 180; 5 values for NVARS (defined as the square value of the number of predictors in the model plus/minus 1 and 2); 4 values for VAR: .000005, .00005, .0005 and .005 and r values for DEPTH: 0 (unrestricted), 15, 20, 25 and 30 . For each of the different values in the grid we estimate the random forest models using half of the sample; we then evaluate performance in the unseen data. Table C.1 presents the hyperparameter values leading to the most accurate predictions (according to the MSE) for each model.

	Preferred Hyperparameters Values			
	NTREES	NVARS	DEPTH	VAR
Model 1	120	3	unrestricted	.000005
Model 2	180	8	20	.0005
Model 3	180	9	unrestricted	.00005
Model 4	180	9	unrestricted	.000005
Model 5	180	8	unrestricted	.000005
Model 6	180	8	unrestricted	.000005

Table C.1. Preferred hyperparameter values employed in Models 1–6.

D. Additional results from re-estimating models in MP13 and MP14 using SED data

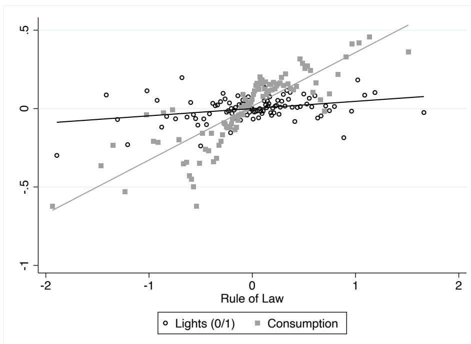
This section presents additional analyses related to Section 7 in the main text. We first present binscatters of the relationships between institutions and nightlights, on one hand, and SED consumption, on the other. These binscatters provide a visual representation of the results presented in Tables 4, 5 and 6 in the main text. We then present estimates of additional models in MP13 and MP14 using SED data.

D.1. Binscatters of institutions and development. The top two graphs in Figure D.1 show binscatters for development and institutions. The solid circles and dark line in graph A show there is no relationship between RULE OF LAW and NIGHTLIGHTS, which is what MP14 find.[1] The graph also shows the strong relationship between RULE OF LAW and CONSUMPTION (the solid squares and light line) that was found in Panel D of Table 4. Note that the two lines show the conditional estimate of $\hat{u}$: at any specific value of RULE OF LAW, the conditional estimate of $\hat{u}$ is the y-value on the gray line minus the y-value on the dark line. This conditional estimate of $\hat{u}$ is positive at low values of RULE OF LAW and becomes negative at high values of RULE OF LAW. Graph B depicts the binscatter for MP13. The value of NIGHTLIGHTS is increasing in JURISDICTIONAL HIERARCHY (as in Panel A of Table 5); there is no relationship between JURISDICTIONAL HIERARCHY and CONSUMPTION (as in Panel D of that table), and the conditional estimate of $\hat{u}$ is increasing as JURISDICTIONAL HIERARCHY increases, suggesting a positive bias in the coefficient estimates of JURISDICTIONAL HIERARCHY in the nightlights models.

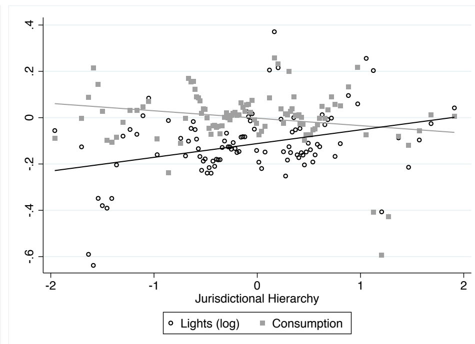
Panel C of Figure D.1 presents the binscatters for lit (diamonds) and dark (squares) pixels separately. Predicted consumption is higher on average in the lit pixels across the range of RULE OF LAW, but within each group of pixels there is a strong relationship between CONSUMPTION and RULE OF LAW, which is completely unaccounted for when nightlights are employed. Panel D shows that within lit and dark pixels, there is no relationship between JURISDICTIONAL

[1] To facilitate interpretation, we have standardized all variables to have a mean of zero and a standard deviation of 1. The binscatters are generated by including controls for log of population density and log of pixel size, as well as indicators for each ethnic group.

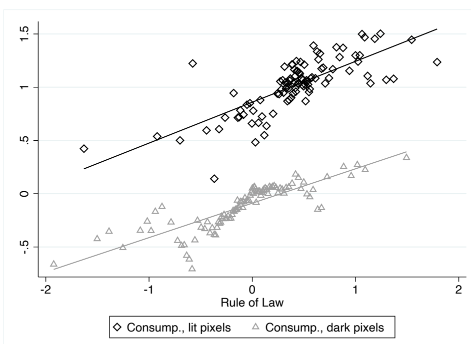
HIERARCHY and consumption, even though JURISDICTIONAL HIERARCHY is positively related to lights.



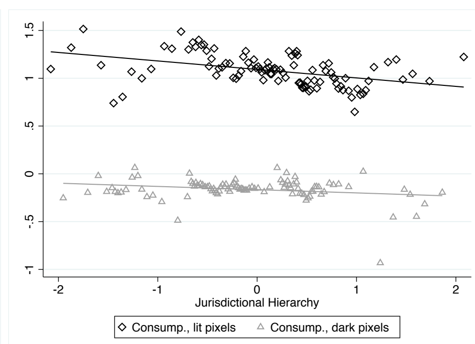
(a)



(b)



(c)



(d)

Figure D.1. NIGHTLIGHTS, CONSUMPTION AND INSTITUTIONS. Panel A shows the binscatters for nightlights and consumption (RF-3) and . Panel B shows the binscatter of nightlights and consumption (RF-3) on jurisdictional hierarchy. Panel C shows the binscatters of consumption (RF-3) on RULE OF LAW in lit and unlit pixels. Panel D shows the binscatters of consumption (RF-3) on JURISDICTIONAL HIEARCHY in lit and unlit pixels. Binscatters include controls for pixel size and population density, as well as ethnicity fixed effects (Panels A and C) or country fixed effects (Panels B and D). All variables are standardized to have mean=0 and SD=1 to facilitate comparisons.

D.2. Estimating additional models in MP13 and MP14. This section presents additional results related to re-analysis of MP13 and MP14 models. As in the main text, we use the MP data and statistical models but substitute cell consumption or cell poverty rate for nightlights. Tables D.1 to D.4 present results from models in MP14. Each table states the models from MP14 that are being re-estimated with the SED data, as well as which SED variable (consumption or poverty) is used as the outcome. The reader is referred the tables in MP14 for model specifics. Together, Tables D.1 to D.4 show a robust relationship between the national institutions variables (RULE OF LAW and CONTROL OF CORRUPTION) and development as measured by the consumption and poverty measures from RF-3. Tables D.5 to D.9 present results from models in MP13. Again, each table describes the model being re-estimated and the dependent variable being used, and the reader can find details about model specifics in MP13. Together, Tables D.5 to D.9 indicate that there is no robust relationship between the ethnic institutions variables and development when SED measures of consumption or poverty from RF-3 are used as the outcome variable.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Dep. variable is poverty, model RF-1				
RULE OF LAW	-0.0323	-0.0139	-0.0289	-0.0145				
	(0.0188)	(0.0085)	(0.0166)	(0.0099)				
CONTROL OF CORRUPTION					-0.0443	-0.0249	-0.0414	-0.0259
					(0.0236)	(0.0139)	(0.0224)	(0.0155)
R-squared	.11	.24	.13	.25	.12	.24	.14	.25
Observations	40871	40871	39250	39250	40871	40871	39250	39250
				Panel B: Dep. variable is poverty, model RF-2				
RULE OF LAW	-0.1538*	-0.0579**	-0.1136*	-0.0437				
	(0.0669)	(0.0221)	(0.0486)	(0.0244)				
CONTROL OF CORRUPTION					-0.2198***	-0.0958***	-0.1705***	-0.0839**
					(0.0637)	(0.0270)	(0.0505)	(0.0301)
R-squared	.16	.73	.32	.74	.25	.73	.38	.75
Observations	40871	40871	39250	39250	40871	40871	39250	39250
				Panel C: Dep. variable is poverty, model RF-3				
RULE OF LAW	-0.1817*	-0.0948**	-0.1433*	-0.0831*				
	(0.0781)	(0.0288)	(0.0619)	(0.0347)				
CONTROL OF CORRUPTION					-0.2647***	-0.1682***	-0.2193***	-0.1579***
					(0.0728)	(0.0363)	(0.0612)	(0.0406)
R-squared	.17	.73	.32	.74	.29	.74	.39	.76
Observations	40871	40871	39250	39250	40871	40871	39250	39250
Ethnicity fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Population density and area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location controls.	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls	No	No	Yes	Yes	No	No	Yes	Yes

Table D.1. RE-ESTIMATING MP14 TABLE 4, PANEL B, USING SED POVERTY RATES. This table re-estimates MP14 Table 4, panel B, using poverty rates from RF models 1-3.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Dep. variable is consumption, model RF-3							
RULE OF LAW	0.4007**	0.2980**	0.3944**	0.2860**				
	(0.1509)	(0.0936)	(0.1478)	(0.0973)				
CONTROL OF CORRUPTION					0.5391***	0.4056***	0.5296***	0.3835**
					(0.1433)	(0.1216)	(0.1415)	(0.1239)
R-squared	.27	.73	.27	.73	.37	.73	.37	.74
Observations	20441	20441	12869	12869	20441	20441	12869	12869
	Panel B: Dep. variable is poverty, model RF-3							
RULE OF LAW	-0.1698*	-0.0955***	-0.1699*	-0.0926***				
	(0.0695)	(0.0276)	(0.0699)	(0.0261)				
CONTROL OF CORRUPTION					-0.2480***	-0.1543***	-0.2466***	-0.1441***
					(0.0700)	(0.0388)	(0.0715)	(0.0369)
R-squared	.17	.7	.18	.71	.29	.71	.29	.72
Observations	20441	20441	12869	12869	20441	20441	12869	12869
Ethnicity fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Population density and area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location controls.	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls	No	No	Yes	Yes	No	No	Yes	Yes

Table D.2. RE-ESTIMATING MP14 TABLE V, PANEL B. This table re-estimates MP14 Table 5, panel B. Panel A uses consumption from RF-3 as the dependent variable and Panel B uses poverty rates from RF-3 as the dependent variable.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Dep. variable is consumption, model RF-3												
RULE OF LAW	0.1901***	0.1901***	0.1868***	0.2028***	0.2001***	0.2135***						
	(0.0596)	(0.0596)	(0.0611)	(0.0634)	(0.0714)	(0.0804)						
	3.19	3.19	3.06	3.20	2.80	2.66						
CONTROL OF CORRUPTION						0.1513**	0.1501**	0.1531**	0.1746**	0.1601**	0.1720**	
							(0.0675)	(0.0654)	(0.0662)	(0.0678)	(0.0760)	(0.0877)
							2.24	2.30	2.31	2.58	2.11	1.96
R-squared	.77	.77	.73	.73	.74	.74	.76	.76	.71	.71	.72	.72
Observations	38438	38438	20441	20441	12869	12869	38438	38438	20441	20441	12869	12869
Panel B: Dep. variable is poverty, model RF-3												
RULE OF LAW	-0.0651***	-0.0651***	-0.0620**	-0.0688***	-0.0690***	-0.0686***						
	(0.0229)	(0.0229)	(0.0250)	(0.0257)	(0.0240)	(0.0263)						
	-2.84	-2.84	-2.48	-2.68	-2.87	-2.61						
CONTROL OF CORRUPTION							-0.0440*	-0.0445*	-0.0469*	-0.0570**	-0.0505*	-0.0480*
							(0.0238)	(0.0230)	(0.0249)	(0.0259)	(0.0259)	(0.0291)
							-1.85	-1.94	-1.89	-2.21	-1.95	-1.65
R-squared	.75	.75	.70	.70	.71	.71	.74	.74	.69	.69	.70	.70
Observations	38438	38438	20441	20441	12869	12869	38438	38438	20441	20441	12869	12869
Ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pixel area and pop. dens.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table D.3. RE-ESTIMATING MP14 TABLE VI. This table re-estimates MP14 Table 6. Panel A uses consumption from RF-3 as the dependent variable and Panel B uses poverty rates from RF-3 as the dependent variable.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Dep. variable is consumption, model RF-3												
RULE OF LAW	0.2187**	0.2022**	0.1974**	0.3181	0.3078***	0.2744***						
	(0.1033)	(0.0855)	(0.0803)	(.)	(0.0929)	(0.0792)						
	2.12	2.36	2.46	.	3.31	3.47						
C.CORRUPT							0.3489***	0.3052***	0.3017***	0.4214***	0.3687***	0.3523***
							(0.0922)	(0.0780)	(0.0780)	(0.0993)	(0.0744)	(0.0694)
							3.78	3.91	3.87	4.24	4.95	5.08
R-squared	.68	.68	.68	.74	.75	.75	.78	.78	.78	.83	.83	.83
Observations	17796	17796	17796	8478	8478	8478	18261	18261	18261	9241	9241	9241
Panel B: Dep. variable is poverty, model RF-3												
RULE OF LAW	-0.0624	-0.0576*	-0.0613*	-0.1079	-0.1081***	-0.1043***						
	(0.0400)	(0.0328)	(0.0325)	(.)	(0.0373)	(0.0327)						
	-1.56	-1.76	-1.89	.	-2.90	-3.19						
C.CORRUPT.							-0.1281***	-0.1072***	-0.1075***	-0.1781	-0.1477***	-0.1429***
							(0.0333)	(0.0274)	(0.0264)	(.)	(0.0381)	(0.0344)
							-3.85	-3.92	-4.08	.	-3.88	-4.15
R-squared	.62	.62	.62	.68	.68	.68	.73	.73	.73	.78	.79	.79
Observations	17796	17796	17796	8478	8478	8478	18261	18261	18261	9241	9241	9241
Ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pixel area and pop. dens.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table D.4. RE-ESTIMATING MP14 TABLE VII. This table re-estimates MP14 Table 6. Panel A uses consumption from RF-3 as the dependent variable and Panel B uses poverty rates from RF-3 as the dependent variable.

	(1)	(2)	(3)	(4)	(5)
	Panel A: Dep. variable is continuous lights				
JURISDICTIONAL HIERARCHY	0.1343	0.1529*	0.1029**	0.1176***	0.0793***
	(0.0969)	(0.0886)	(0.0450)	(0.0454)	(0.0305)
Adjusted R-squared	0.01	0.19	0.28	0.32	0.33
Observations	61359	61359	61359	61015	61015
	Panel B: Dep. variable is poverty				
JURISDICTIONAL HIERARCHY	0.0054	-0.0036	0.0012	-0.0012	-0.0012
	(0.0245)	(0.0065)	(0.0080)	(0.0078)	(0.0074)
Adjusted R-squared	0.00	0.7	0.72	0.76	0.77
Observations	61359	61359	61359	61015	61015
Country Fixed effects	No	Yes	Yes	Yes	Yes
Population Density	No	No	Yes	Yes	Yes
Controls at the Pixel level	No	No	No	Yes	Yes
Controls at the Ethnic-Country level	No	No	No	No	Yes

Table D.5. RE-ESTIMATING MP13, TABLE V PANEL A. This table re-estimates MP13 Table V, Panel A. Panel A re-estimates models 6-10 using the MP13 measure of log of nightlights. Panel B re-estimates models 1-5 using the cell poverty rate as the dependent variable.

	(1)	(2)	(3)	(4)	(5)
	Panel A: Dep. variable is consumption				
Petty Chiefdoms	-0.1587*	0.0193	0.0077	-0.0040	0.0002
Double-clustered s.e.	(0.0828)	(0.0289)	(0.0270)	(0.0268)	(0.0280)
Paramount Chiefdoms	-0.1201	-0.0200	-0.0366	-0.0252	-0.0177
Double-clustered s.e.	(0.1100)	(0.0331)	(0.0302)	(0.0216)	(0.0208)
Pre-ColonialStates	-0.0538	-0.0032	-0.0273	-0.0312	-0.0342
Double-clustered s.e.	(0.1276)	(0.0460)	(0.0535)	(0.0447)	(0.0469)
Adjusted R-squared	0.02	0.78	0.79	0.82	0.83
	Panel B: Dep. variable is poverty				
Petty Chiefdoms	0.0937*	-0.0114	-0.0041	0.0015	-0.0007
Double-clustered s.e.	(0.0545)	(0.0160)	(0.0147)	(0.0137)	(0.0152)
Paramount Chiefdoms	0.0766	0.0017	0.0120	0.0030	-0.0022
Double-clustered s.e.	(0.0649)	(0.0142)	(0.0137)	(0.0157)	(0.0152)
Pre-Colonial States	0.0128	-0.0229	-0.0078	-0.0063	-0.0033
Double-clustered s.e.	(0.0741)	(0.0242)	(0.0285)	(0.0239)	(0.0250)
Adjusted R-squared	0.02	0.7	0.72	0.76	0.77
Country Fixed effects	No	Yes	Yes	Yes	Yes
Population Density	No	No	Yes	Yes	Yes
Controls at the Pixel level	No	No	No	Yes	Yes
Controls at the Ethnic-Country level	No	No	No	No	Yes
Observations	61359	61359	61359	61015	61015

Table D.6. Re-estimating MP13, Table V panel B. This table re-estimates MP13 Table V, Panel B. Panel A re-estiamtes models 1-5 using consumption from RF-3 as the dependent variable. Panel B re-estimates these models 1-5 using the cell poverty rate from RF-3 as the dependent variable.

	(1)	(2)	(3)
	Panel A: Dep. variable is consumption		
Jurisdictional Hierarchy	-0.0116	-0.0151	-0.0138
Double-clustered s.e.	(0.0134)	(0.0155)	(0.0139)
Adjusted R-squared	0.83	0.85	0.86
	Panel B: Dep. variable is poverty		
Jurisdictional Hierarchy	0.0029	0.0051	0.0040
Double-clustered s.e.	(0.0076)	(0.0086)	(0.0082)
Adjusted R-squared	0.73	0.76	0.77
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes
Population Density	No	Yes	Yes
Controls at the Pixel level	No	No	Yes
Observations	72545	72545	72297

Table D.7. Re-estimating MP13, Table VII. This table re-estimates the first three models (which use all observations) from MP13 Table VII. The dependent variable in Panel A is consumption from RF-3. The dependent variable in Panel B is the cell poverty rate from RF-3.

	Dependent Variable: Consumption		
	< 100 km of ethnic border (1)	< 150 km of ethnic border (2)	< 200 km of ethnic border (3)
Panel A: Pre-Colonial Ethnic Institutions and Regional Development Within Contiguous Ethnic Homelands in the Same Country Pixel-Level Analysis in Areas Close to the Ethnic Border			
Panel 1: Border Thickness—Total 50 km (25 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	0.0117	0.0083	0.0090
Double-clustered s.e.	(0.0099)	(0.0090)	(0.0082)
	1.18	0.92	1.10
Adjusted R-squared	.89	.89	.89
Observations	6237	9476	11920
Panel 2: Border Thickness—Total 100 km (50 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	0.0092	0.0052	0.0074
Double-clustered s.e.	(0.0098)	(0.0088)	(0.0081)
	0.94	0.59	0.92
Adjusted R-squared	.89	.89	.89
Observations	4053	7292	9736
Panel B: Pre-Colonial Ethnic Institutions and Regional Development Within Contiguous Ethnic Homelands in the Same Country Pixel-Level Analysis in Areas Close to the "Thick" Ethnic Border Border Controlling for a Fourth-order RD-Type Polynomial in Distance to the Ethnic Border			
Panel 1: Border Thickness—Total 50 km (25 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	0.0134	0.0045	0.0064
Double-clustered s.e.	(0.0176)	(0.0165)	(0.0160)
	0.76	0.27	0.40
Adjusted R-squared	.89	.89	.89
Observations	6237	9476	11920
Panel 2: Border Thickness—Total 100 km (50 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	0.0191	0.0029	0.0053
Double-clustered s.e.	(0.0207)	(0.0180)	(0.0161)
	0.92	0.16	0.33
Adjusted R-squared	.89	.89	.89
Observations	4053	7292	9736
RD-Type Polynomial	Yes	Yes	Yes
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes
Population Density	Yes	Yes	Yes
Controls at the Pixel level	Yes	Yes	Yes

Table D.8. Re-estimating MP13, Table VIII using consumption. This table re-estimates the first three models (which use all observations) from MP13 Table VIII. The dependent variable is consumption from RF-3.

	Dependent Variable: Poverty		
	< 100 km of ethnic border (1)	< 150 km of ethnic border (2)	< 200 km of ethnic border (3)
Panel A: Pre-Colonial Ethnic Institutions and Regional Development Within Contiguous Ethnic Homelands in the Same Country Pixel-Level Analysis in Areas Close to the Ethnic Border			
Panel 1: Border Thickness—Total 50 km (25 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	-0.0054	-0.0032	-0.0040
Double-clustered s.e.	(0.0055)	(0.0051)	(0.0050)
	-0.97	-0.61	-0.80
Adjusted R-squared	.83	.82	.82
Observations	6237	9476	11920
Panel 2: Border Thickness—Total 100 km (50 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	-0.0034	-0.0009	-0.0026
Double-clustered s.e.	(0.0061)	(0.0055)	(0.0054)
	-0.56	-0.16	-0.49
Adjusted R-squared	.83	.82	.82
Observations	4053	7292	9736
Panel B: Pre-Colonial Ethnic Institutions and Regional Development Within Contiguous Ethnic Homelands in the Same Country Pixel-Level Analysis in Areas Close to the "Thick" Ethnic Border Border Controlling for a Fourth-order RD-Type Polynomial in Distance to the Ethnic Border			
Panel 1: Border Thickness—Total 50 km (25 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	-0.0030	0.0061	0.0018
Double-clustered s.e.	(0.0107)	(0.0091)	(0.0083)
	-0.28	0.68	0.21
Adjusted R-squared	.83	.82	.82
Observations	4053	7292	9736
Panel 2: Border Thickness—Total 100 km (50 km from each side of the ethnic boundary)			
Jurisdictional Hierarchy	-0.0030	0.0061	0.0018
Double-clustered s.e.	(0.0107)	(0.0091)	(0.0083)
	-0.28	0.68	0.21
Adjusted R-squared	.83	.82	.82
Observations	4053	7292	9736
RD-Type Polynomial	Yes	Yes	Yes
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes
Population Density	Yes	Yes	Yes
Controls at the Pixel level	Yes	Yes	Yes

Table D.9. Re-estimating MP13, Table VIII using poverty. This table re-estimates the first three models (which use all observations) from MP13 Table VIII. The dependent variable is the poverty rate from RF-3.

E. References

Berry, P. A. M., Smith, R., and Benveniste, J., (2010), "ACE2: The New Global Digital Elevation Model" In Mertikas, S. (eds) Gravity, Geoid and Earth Observation, International Association of Geodesy Symposia, Berlin, Heidelberg, Springer.

Cagé, Julia, and Valeria Rueda,(2016) "The Long-Term Effects of thePrinting Press in sub-Saharan Africa," *American Economic Journal: Applied Economics* 8, 69–99.

Cage, Julia, and Valeria Rueda, (2020), "Sex and the Mission: The Conflicting Effects of Early Christian Investments on the HIV Epidemic in Sub-Saharan Africa," *Journal of Demographic Economics* , Volume 86 ,(3). 213 - 257.

Doxsey-Whitfield, Erin, Kytt MacManus, Susana B. Adamo, Linda Pistolesi, John Squires, Olena Borkovska and Sandra R. Baptista (2015), "Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4," *Papers in Applied Geography*, 1:3, 226-234, DOI: 10.1080/23754931.2015.1014272.

EDGAR (Emissions Database for Global Atmospheric Research) Community GHG Database (a collaboration between the European Commission, Joint Research Centre (JRC), the International Energy Agency (IEA), and comprising IEA-EDGAR CO2, EDGAR CH4, EDGAR N2O, EDGAR F-GASES version 7.0, (2022) European Commission, JRC (Datasets).

Elvidge, Christopher D., Baugh, Kimberly E., Dietz, John B., Bland, Theodore, Sutton, Paul C., and Kroehl, Herbert W., (1999), "Radiance Calibration of DMSP-OLS Low-Light Imaging Data of Human Settlements," *Remote Sensing of Environment*, 68, 77–88.

Elvidge, Christopher D, K. Baugh, M. Zhizhin, F. C. Hsu, and T. Ghosh, (2017), "VIIRS night-time lights," *International Journal of Remote Sensing*, vol. 38, pp. 5860–5879.

Elvidge, Christopher D., Baugh, Kimberly E., Zhizhin, Mikhail, and Hsu, Feng-Chi, (2013), “Why VIIRS data are superior to DMSP for mapping nighttime lights,” *Proceedings of the Asia-Pacific Advanced Network*, 35 (2013), 62.

Galor, Oded, and Zak, Mer, (2015) “Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability,” *Agricultural Suitability*, July 12, 2015.

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

Harris, Ian, Osborn, Timothy J., Jones, Phil, and Lister, David, (2020), “Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset,// *Scientific Data*, 7, 109.

Hsu, F. C., Baugh, K. E., Ghosh, T., Zhizhin, M., Elvidge, C. D. (2015), “DMSP-OLS radiance calibrated nighttime lights time series with intercalibration,” *Remote Sensing*, 7(2), 1855-1876.

Lehner, Bernhard, and Petra Döll, (2004), “Development and validation of a global database of lakes, reservoirs and wetlands,” *Journal of hydrology*, 296, 1–22.

Lujala, Päivi; Jan Ketil Rød and Nadia Thieme, (2007), “Fighting over Oil: Introducing A New Dataset,” *Conflict Management and Peace Science* 24(3), 239-256.

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

Michalopoulos, S., and E. Papaioannou (2013), “Pre-Colonial Ethnic Institutions and Contemporary African Development,” *Econometrica* 81, 113–152.

Michalopoulos, S., and E. Papaioannou (2014), “National Institutions and Subnational Development in Africa,” *The Quarterly Journal of Economics* 129, 151–213.

Nunn, Nathan, (2010), “Religious Conversion in Colonial Africa,” *American Economic Review*, 100 (2010), 147–152.

Nunn, Nathan, and Puga, Diego, “Ruggedness: The Blessing of Bad Geography in Africa,” *Review of Economics and Statistics* 94 (2012), 20–36.

Olson, David M, and Dinerstein, Eric, (2002), “The Global 200: Priority ecoregions for global conservation,” *Annals of the Missouri Botanical Garden*,” 199–224.

Weiss, Daniel J, Lucas, Tim C D, Nguyen, Michele, Nandi, Anita K, Bisanzio, Donal, Battle, Katherine E, Cameron, Ewan, Twohig, Katherine A, Pfeffer, Daniel A, Rozier, Jennifer A, Gibson, Harry S, Rao, Puja C, Casey, Daniel, Bertozzi-Villa, Amelia, Collins, Emma L, Dalrymple, Ursula, Gray, Naomi, Harris, Joseph R, Howes, Rosalind E, Kang, Sun Yun, Keddie, Suzanne H, May, Daniel, Rumisha, Susan, Thorn, Michael P, Barber, Ryan, Fullman, Nancy, Huynh, Chantal K, Kulikoff, Xie, Kutz, Michael J, Lopez, Alan D, Mokdad, Ali H, Naghavi, Mohsen, Nguyen, Grant, Shackelford, Katya Anne, Vos, Theo, Wang, Haidong, Smith, David L, Lim, Stephen S, Murray, Christopher J L, Bhatt, Samir, Hay, Simon I, and Gething, Peter W, (2019), “Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000–17: a spatial and temporal modelling study,” *The Lancet*, 394 (2019), 322–331.

Yeh, C., A. Perez, A. Driscoll, G. Azzari, Z. Tang, D. Lobell, S. Ermon, and M. Burke (2020) “Using publicly available satellite imagery and deep learning to understand economic well-being in Africa,” *Nature Communications*, 11, 2583.