Appendix (For Online Publication)

This appendix is divided into six sections:

- Section A shows a figure documenting the increase in the use of nightlights in economics over time.
- Section B lists the DHS surveys used in constructing the training variable, and presents further details about its construction.
- Section C presents a table that describes the variables used in the prediction exercise.
- Section D provides additional information about the prediction models, including parameter tuning, variable importance in the prediction models, prediction accuracy, and a comparison of the prediction results from our RF models with those of Yeh et al. (2020) (section D.5).
- Section E provides further information about the approach we follow to estimate celllevel poverty rates.
- Section F provides additional information and estimation results regarding the analysis of the models in Michalopoulos and Papaioannou (2013 and 2014) using the SED data.

A. Nightlights in Economic Research

Figure A.1 displays the evolution of the number of papers referencing "nightlights" in economics according to Google Scholar. For additional information about the use of nightlights over time in economics research, see Gibson, et al (2020).



Figure A.1. PAPERS IN ECONOMICS REFERENCING NIGHTLIGHTS. Source: Google Scholar. The graph depicts the number of papers in Google Scholar obtained using the keywords "nighlights+economics" from 2005 to 2022.

B. Constructing the training variable, $\widehat{y_{rct}^*}$.

This section provides additional details about the construction of the training variable, $\widehat{y_{rct}^*}$. Here we summarize the main steps involved in the process; the subsequent sections provide additional details related to each of these steps.

- (1) For each survey, estimate a principal components model using all respondents and a variety of asset variables that are present in the given survey to generate the asset measure, y_{ict}^{A} .
- (2) Use the WB-PIP data on average consumption and the Gini coefficient for the countryyear of the survey to obtain estimates of μ_{ct}^C and σ_{ct}^C , the mean and the standard deviation of log-consumption measured in 2011 PPP dollars.
- (3) Apply the estimates of y_{ict}^A , μ_{ct}^C and σ_{ct}^C to compute $\widehat{y_{ict}^*}$ as described in section 3. The resulting $\widehat{y_{ict}^*}$ is expressed in log of consumption in 2011 PPP dollars.
- (4) Calculate $\widehat{y_{rct}^*}$ by averaging $\widehat{y_{ict}^*}$ for all *i* in cluster *r*.

B.1. Constructing the asset index, y_{ict}^A . We use 85 DHS surveys comprising over 900,000 households who are sampled from across 29 sub-Saharan African countries in the period 2006-2018. Figure B.1 displays a map of the African countries for which we have DHS data, and Table B.1 provides a list of all the DHS surveys used to create the training data. DHS surveys provide information on household-level ownership of different assets. These asset variables are related to sanitation in the home (the source of drinking water and type of toilet facility), the nature of the household's dwelling (flooring, wall, and roof materials; presence of electricity and number of sleeping rooms), and the presence of particular assets (e.g., radio, television, refrigerator, motorcycle or scooter, car or truck, telephone and mobile phone).

For each individual survey, we use principal component analysis (PCA) to create an asset index. The (unrotated) loadings of the asset variables on the first component are used to predict an aggregate asset score for each household. Since we estimate the PCA separately for each survey, the loadings vary across surveys. That is, we do not assume that the relationship of each asset to consumption is the same across time and space, which provides additional flexibility in constructing the index. Since DHS data is defined at the household level, we follow Duclos et al (2004) and divide the resulting index by $s^{0.5}$ to obtain a per-capita index, where *s* is the size of the household.¹ Finally, to calculate y_{ict}^A , we take the log of the index and standardize it so that for each survey the index has a mean of 0 and a standard deviation of 1.

B.2. Estimates of $\mu_{y_{ct}^C}$ and $\sigma_{y_{ct}^C}$. WB-PIP data, which are based on surveys rather than national accounts, provide the current gold standard for country level estimates of consumption in sub-Saharan Africa. WB-PIP provides estimates of the mean of consumption per capita but not of the mean and variance of its log, $\sigma_{y_{ct}^C}$ or $\mu_{y_{ct}^C}$. To overcome this limitation, we assume that y_{ict}^C follows a normal distribution with mean $\mu_{y_{ct}^*}$ and variance $\sigma_{y_{ct}^2}^2 = \sigma_{y_{ct}^*}^2 + \sigma_{e_{ct}}^2$, (see equation (3)). This implies that consumption per capita (the exponential value of y_{ict} , which we denote by x_{ict}) follows a log-normal distribution.² The log-normality assumption makes it possible to use estimates of the Gini coefficient, which exist in WB-PIP country-level data (as opposed to estimates of the variance of the distribution, which do not), to estimate the

¹All the analysis in the paper has also been done without applying this correction and the results are virtually identical.

²See Battistin, Blundell and Lewbel (2009) for recent evidence supporting this claim.



Figure B.1. Countries in the training sample

standard deviation of log consumption. Under log-normality, the Gini coefficient associated with the consumption level, x_{ct}^C , is related to the standard deviation of the log of consumption, y_{ct}^C as follows:

$$GINI_{x_{ct}^{C}} = 2\Phi(\frac{\sigma_{y_{ct}^{C}}}{\sqrt{2}}) - 1, \qquad (b.1)$$

where $\Phi(.)$ is the cumulative standard normal distribution. Therefore, the log-normality assumption allows us to estimate the variance of log-consumption, $\sigma_{y_{ct}^c}^2$, from estimates of the Gini of total consumption, which are typically available in WB-PIP. The log-normality assumption of the WB-PIP data also makes it straightforward to estimate $\mu_{y_{ct}^c}$ because the mean of total consumption, x_{ct}^C , and its log are related through the following equation:

$$\mu_{x_{ct}^{\rm C}} = e^{\left(\mu_{y_{ct}^{\rm C}} + 0.5(\sigma_{y_{ct}^{\rm C}}^2)\right)}_{3}$$

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

Finally, we apply the estimates of y_{ict}^A , μ_{ct}^C and σ_{ct}^C to compute $\widehat{y_{ict}^*}$ as described in section 3.1. The resulting $\widehat{y_{ict}^*}$ is expressed in log of consumption in 2011 PPP dollars.

A central challenge when using asset variables to compute a measure of economic wellbeing is achieving comparability across time and space. It is useful to contrast the way this comparability is achieved when using $\widehat{y_{ict}^*}$ with the approach in Yeh et al. (2020).³ Yeh et al. (2020) achieve comparability in their asset-based variable by pooling all DHS surveys,

³Chi et al. (2022) also compute a training variable based on a DHS asset index, but this training variable only provides within country-year information, i.e., its values cannot be compared across countries or over time as all survey-years are standardized to have a zero mean.

estimating a principal component model using a set of asset variables that is commonly available in each survey, and using the factor scores to generate the measure of respondent's economic well-being. This results in a variable that though comparable across time and space, captures only ordinal differences in economic well-being. Thus, like nightlights, the variable has no substantively meaningful metric.

The most important difference between the Yeh et al approach and the approach here is of course that $\widehat{y_{ict}^*}$ is expressed as consumption per capita in 2011 PPP dollars. But there are other differences worth underscoring as well. First, the pooled PCA approach in Yeh et al requires a common set of assets in every survey. This limits the nature of the training data: since all surveys must contain the same set of asset variables, surveys must be dropped when they do not include the requisite asset variables. One way to limit the problem is to use a relatively small number of asset variables in constructing the measure of well-being. But this strategy limits the variability in the measure. By allowing the nature of the asset variables used in constructing $\widehat{y_{ict}^*}$ to vary across surveys, the approach here avoids both of these limitations. Second, the pooled PCA achieves comparability by assuming that the relationship between assets and economic well-being is the same across countries and over time. This is a strong assumption. We might expect, for example, that the relationship between owning a bicycle (or radio or computer or cell phone) and economic well-being to vary across countries or over time. The approach here avoids this strong assumption. Instead, it achieves comparability through the use of macro data on consumption and inequality. While this approach may have the problems discussed in section 3.2, the fact that $\widehat{y_{ict}^*}$ is denominated in dollars – and that it can be used to derive poverty rates - opens avenues for evaluating the measure that are unavailable when using measures lacking an interpretable metric.

B.3. From individuals to clusters. The final step is to average $\widehat{y_{ict}^*}$ to the enumeration area level, also called "cluster," which is roughly equivalent to villages in rural areas or neighborhoods in urban areas, as this is the level at which geo-coordinates are available in the public survey data. Since $\widehat{y_{ict}^*}$ is expressed in logs, we exponentiate it to obtain consumption per capita, compute the cluster-level average and then take the log of the average to obtain $\widehat{y_{rct}^*}$.

the log of cluster mean consumption in 2011 PPP dollars for cluster r, c, t.⁴ For each cluster, DHS publishes the latitude and longitude of the cluster's "centroid." For privacy reasons, DHS randomly jiggers the published location of this centroid by up to 5 km from the true centroid. We therefore assign each DHS cluster to a 10x10 square kilometer pixel that has the centroid reported by DHS (thereby ensuring that the square encompasses the true centroid).⁵ We denote this geo-located variable as $\widehat{y_{rct}^*}$, where *r* denotes enumeration area (cluster). The training data we use in the prediction models include 34,484 clusters.

⁴We drop clusters with invalid GPS coordinates. To limit measurement error due to small numbers of households in a cluster, we also drop clusters with less than 16 households (which results in dropping 2% of the clusters). On average, there are 26 respondents per cluster.

⁵Yeh et al (2020) estimate that the jiggering by DHS degrades model performance, reducing the R^2 by about 0.07.

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-vector
d_diamonds	Distance in meters from cell centroid to nearest diamond deposit. See Gilmore et al (2005).	https://www.prio.org/data/10
	Table conti	nues on next page.

C. Table describing predictors and their sources

	Table des	cribing 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)

PBENCTOP		uig preuctors and liter sources, continued Compete
	DEFINITION	JOURLE
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)
o 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 describing predictors and their sources, continued

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)

contin
sources,
their
and
predictors
lescribing 1
<u> </u>

D. Additional information about the predictions models and their accuracy.

This section provides information about parameter tuning (section D.1) and about the variables with the largest importance in the prediction models (section D.2). It also presents additional information about prediction accuracy, including an analysis related to the poor performance of the models including only nightlights as predictors (section D.3). Finally, it compares the prediction results from our RF models with those of Yeh et al. (2020) (section D.5).

D.1. Parameter tuning. The main hyperparameters of the RF 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. 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. Using this process, we identify the hyperparameter values leading to the lowest MSE for each model. Table D.1 presents the resulting hyperparameters employed in the three models.

	F	referred	l Hyperj	parame	ters Values
	NTREES	NVARS	DEPTH	VAR	MIN OBS PER LEAF
Model 1	180	1	25	.0001	7
Model 2	180	8	35	.0001	3
Model 3	180	6	35	.0005	1

 $\label{eq:table_$

D.2. Variable importance. Table D.2 shows the most informative predictors in models RF-2 and RF-3. The most important variable in both models is whether a cell is located in the desert. Nightlights variables are important in RF-2, and CO2 emissions, population density, the disease environment, and variables related to a cell's location are important in both models.

D.3. Further evaluation of prediction accuracy. Figure D.1 displays the MSE and the R² from RF-1 through RF-3 using all 85 out-of-sample sets of forecasts. Both panels show that prediction accuracy varies across surveys. But this is especially true for RF-1. Focusing on panel (a), RF-1 generates a very large MSE for some surveys and substantial dispersion of

	Relative Varial	ble Importance
Ranking	RF-2	RF-3
1	Desert ecosystem (1)	Desert ecosystem (1)
2	NLs (3 yr mean) (.34)	CO ₂ (.25)
3	CO ₂ (.27)	Population Density (.16)
4	Population Density (.16)	Latitude (.08)
5	Latitude (.14)	Grassland ecosystem (.08)
6	Grassland ecosystem (.11)	Longitude (.07)
7	NL(VIIRS) (.10)	Remoteness (.07)
8	Longitude (.09)	Disease (.06)
9	Disease (.09)	Malaria Incidence (.05)
10	NL(DSMP, blur) (.09)	Grassland ecosystem (.05)

Table D.2. VARIABLE IMPORTANCE. This table provides the 10 most important predictors for models RF-2 and RF-3, together with their relative importance. Importance is relative to the most informative one (whose importance is normalized to 1).

the MSE across the surveys. In RF-2 and RF-3, by contrast, there are a half-dozen surveys that have especially poor performance (though much better than the MSE in RF-1), and the remaining MSEs are concentrated in quite low values, especially in our main model, RF-2.



Figure D.1. OUT-OF-SAMPLE PREDICTION ACCURACY. This figure provides the MSE and \mathbb{R}^2 for the 85 out-of-sample sets of predictions, corresponding to each of the surveys in our sample. Box and Whisker plots are displayed in red.

D.4. Understanding the poor performance of NL-only models. To understand the poor performance of the nightlights model, Figure D.2 depicts the squared correlation between the out of sample predictions and the training data (R^2) from RF-1 (NL only) and RF-2 (the full model) when increasing the number of deciles of data, *X*, used in estimation. For X=2, for instance, the graph depicts the value of the R^2 obtained when only the first two deciles



Figure D.2. PERFORMANCE FOR INCREASING SHARES OF DATA USED IN ESTIMATION. The figure plots the R^2 s from models estimated on the X smallest deciles of the training data. E.g., if X=2, estimation is carried out on the first 2 deciles of the data.

of the training variable are used.⁶ The graph shows that RF-1 forecast accuracy is basically zero when as much as the first six deciles are used for estimation, and it remains quite low until 90% of the observations are employed. The graph therefore confirms that nightlights alone have no power to predict variation in economic well-being for almost 60% of the data. What nightlights make possible is to distinguish the 90% of poorest clusters from the 10% of richest ones. This is unsurprising given the vast areas of populated darkness described in the Introduction. By contrast, the performance of RF-3 is is stronger across the deciles, with an R² around .5 for these first five deciles, which then grows by over 20% when the remaining deciles are included in the estimation sample. This graph therefore highlights that in comparison to the NL-only models, the models with a richer set of predictors not only have power to distinguish the poor from the rich, but also to distinguish the poor from the very poor.

D.5. Comparison with previous benchmarks: Yeh et al. (2020). Yeh et al (2020) present an innovative method based on combining nightlights and daytime imagery to predict an asset wealth index, which is computed using 43 DHS surveys across almost 20,000 African clusters. They train a convolutional neural network (CNN) to predict the cluster-specific

⁶Out-of-sample predictions are obtained as follows: (1) the clusters in the training data are divided into ten "decile" data sets, each including the clusters in the decile and all lower deciles; (2) for each decile data set, RF-1 and RF-2 are estimated, but omitting the data from a held out survey; (3) out-of-sample predictions are obtained for the held-out clusters in each decile data set; (4) the R² for each decile data set is obtained using the out-of-sample predictions.

measure of wealth using temporally and spatially matched multispectral daytime imagery and nightlights as inputs. Based on prediction performance on held out locations, they show that (1) the median squared correlation coefficient between their training variable and the (out-of-sample) predictions is 70% in their best models, and (2) a simple K-nearest neighbor (KNN) model whose only predictor is nightlights has similar performance to a CNN with both nightlights and (the much heavier) daylight imagery.

Our framework is similar to that of Yeh et al. (2020), but there are several differences: (1) they use a different training variable (a unitless index of asset wealth, computed as the first principal component in a sample that pools all DHS surveys), (2)) their sample is much smaller (43 surveys versus 85, in our case), and (3) their country composition is different. Consequently, the results presented in section 4.3 are not directly comparable to the results in Yeh et al. To investigate the relative performance of both approaches, in this section we use random forest models with a large number of variables to predict the training variable from the Yeh et al data set. This allows us to attribute any performance differences to the different methods and predictors we employ.

	MSE and R ² (media	n value))
	MSE	R ²
Yeh, KNN	.191	.691
Yeh, CNN	.179	.687
RANDOM FOREST, WITH RICH SET OF PREDICTORS	.168	.724

Table D.3. COMPARING PREDICTIONS RESULTS FROM YEH ET AL (2020) WITH THOSE FROM RANDOM FOREST MODELS USING THE YEH ET AL TRAINING DATA. Table provides model results when using the training variable from the Yeh et al data set. TheYEH, KNN and YEH, CNN results are obtained from the Yeh et al replication materials.

For the results from Yeh et al, we focus on (a) their model with best performance, a CNN that uses both NLs and daylight imagery as inputs, and (b) their KNN model that includes only NLs (as the latter is much less computationally intensive, and Yeh et al. conclude that its performance is very similar to that of more complex models). We compare performance to a random forest model that includes a rich set of cell-level predictors, as we do in RF-2. All results are based on held-out locations, as described in the main text. Table D.3 presents the results.

Two conclusions stand out. First, when the preferred MSE metric is employed, the performance of the CNN model is significantly better than that of the simpler KNN model (an improvement of 6% in the median MSE). Second, the random forest model with a rich set of predictors outperform those in Yeh et al. (2020), producing a MSE that is 12% smaller than that of the KNN model and 6% smaller than that of the CNN model. Importantly, these improved results are achieved at a much lower computational cost. The use of daytime imagery together with the CNN algorithm is a very expensive computational approach. This complexity and computational expense make it extremely challenging to scale up the predictions to compute maps for the whole of Africa over time. By contrast, the random forest models with a large number of predictors can easily be run in STATA on a personal computer.

E. Estimating cell-level poverty rates

This section provides further details about the nonparametric approach we propose for estimating cell poverty rates. We must first classify clusters into *K* groups, and then estimate the poverty rate of clusters belonging to group *k*, for k = 1...K. We do this in three steps:

- (1) We assume K = 100. To allocate clusters to groups we could use clustering methods, such as k-means. However, we will assume that all clusters in group k have nearly identical average consumption, making it natural to use consumption to assign clusters to groups. Therefore, groups are defined as percentiles of the distribution of $\widehat{y_{rct}}$. We use the percentile values of each group to identify the cut-points dividing a group from its adjacent groups. For example, the 75th group includes all clusters that have a value of $\widehat{y_{rct}}$ between 7.194 and 7.227. Like group 75, the range of cluster mean consumption in each group is very narrow, with an average range that is 0.031.
- (2) We assign each of the roughly 920,000 respondents in the DHS surveys to the group associated with the respondent's cluster mean income. Since the 75th group includes all clusters that have a value of $\widehat{y_{rct}^*}$ between 7.194 and 7.227, if a DHS survey respondent resides in a cluster, for example, with $\widehat{y_{rct}^*} = 7.21$, the respondent would be placed in the 75th group, along with all other respondents across all surveys who reside in a cluster

with a value of $\widehat{y_{rct}^*}$ between 7.194 and 7.227. This particular group has 8,790 individual DHS respondents, which is a typical number of survey respondents in each group.

(3) We use all the DHS respondents in a group (such as the 8,790 in group 75) to calculate the poverty line for the group. Our focus will be on the \$1.90 per day used with 2011 PPP dollars. Using the 8,790 estimates of $\widehat{y_{rct}^*}$ in group 75, we find that 19.1% of the households are below the poverty line of \$1.90 per day using 2011 PPP dollars. Thus, the poverty rate for group 75 is 19.1% and all clusters assigned to this group are assigned this same poverty rate.

Panel (a) in Figure E.1 shows the histogram of cluster-level poverty rates based on \hat{y}_{rct}^* . The blue histogram uses the naive approach, and the stark binary nature of the distribution is what motivates the nonparametric method. The red histogram uses the nonparametric approach to generate the poverty rates, and though there is still slightly more mass at the tails of the distribution, there is a relatively even distribution across the range of poverty rates.



Figure E.1. The DISTRIBUTION OF CLUSTER-LEVEL POVERTY RATES USING NAIVE AND NON-PARAMETRIC APPROACHES. Panel (a) depicts the histogram of poverty rates in DHS clusters using the naive and non-parametric approaches. The poverty line is \$1.90 a day using 2011 PPP dollars. Panel (b) presents scatterplots of log consumption versus (transformed) asset-based indices, along with the 45-degree line.

By aggregating these rates to the country level using using DHS weights, we can gain further insights into how the two approaches differ. We have calculated the national level poverty rates based on the individual-level data, $\widehat{y_{ict}^*}$. Panel (b) in Figure E.1 compares these poverty rates (the y-axis) with the poverty rates based on applying the naive and nonparametric approaches to the cluster-level data. When aggregated to the national level, both approaches use the cluster-level data to produce national poverty rates that are close to those from the individual-level data. But the naive approach tends to underestimate poverty at the higher end of the poverty distribution, and especially, to overestimate poverty at the lower end of the poverty distribution. The same patterns do not exist for the non-parametric estimates. As noted, the goal of the non-parametric approach is to produce fine-grained estimates of poverty rates for cell-level analyses. The evidence here suggests that by doing so, we also replicate national poverty rates more accurately than does the naive approach.

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

This section presents additional results related to re-analysis of MP13 and MP14 models.

Additional MP14 models. Tables F.1 to F.3 present results from models in MP14. Each table states the models from MP14 that are being re-estimated, and the reader is referred the tables in MP14 for model specifics. Together, Tables F.1 to F.3 show that there is no robust relationship between national institutions and nightlights, as in MP14. By contrast, the results for RULE OF LAW are positive and precisely estimated in every model, and the results for CONTROL OF CORRUPTION are precisely estimated in all models except the spatial regression discontinuity models in MP 14 Table VI.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Par	nel A: Dep.	variable is l	NL		
RULE OF LAW	0.0994***	0.0168	0.1031**	0.0159				
	(0.0384)	(0.0181)	(0.0407)	(0.0162)				
CONTROL OF CORRUPTION					0.1292***	0.0269	0.1346***	0.0197
					(0.0471)	(0.0259)	(0.0492)	(0.0210)
Adj. R-squared	0.134	0.319	0.137	0.345	0.144	0.319	0.149	0.345
N	21289	21289	13408	13408	21289	21289	13408	13408
			Panel B: De	ep. variable	is consum	otion, RF-2		
RULE OF LAW	0.4474***	0.2382***	0.4365***	0.2139***				
	(0.1677)	(0.0800)	(0.1590)	(0.0768)				
CONTROL OF CORRUPTION					0.5823***	0.2794***	0.5650***	0.2367**
					(0.1598)	(0.1017)	(0.1530)	(0.0982)
Adj. R-squared	0.227	0.773	0.230	0.784	0.295	0.770	0.296	0.780
N	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 F.1. Re-estimating MP14 Table V, panel B. This table re-estimates MP14 Table 5, panel B.

	(1)	(2)	(3)	(4)	(5)	(9)	(2)	(8)	(6)	(10)	(11)	(12)
			Pane	el A: Dep. v	ariable is N	١L						
RULE OF LAW (HIGH)	0.0166 (0.0138)	0.0058 (0.0135)	0.0123 (0.0165)	0.0050 (0.0185)	-0.0010 (0.0239)	0.0122 (0.0271)						
CONTROL OF CORRUPTION (HIGH)							0.0038 (0.0154)	-0.0086 (0.0137)	0.0116 (0.0166)	0.0123 (0.0176)	-0.0030 (0.0225)	0.0079 (0.0278)
Adj. R-squared N	0.342 40209	0.342 40209	0.320 21289	0.320 21289	0.347 13408	0.347 13408	0.342 40209	0.343 40209	0.320 21289	0.320 21289	0.347 13408	0.347 13408
		P	anel B: De	p. variable	is consump	tion, RF-2						
RULE OF LAW (HIGH)	0.1268*** (0.0415)	0.1292*** (0.0384)	0.0920** (0.0389)	0.1069*** (0.0384)	0.1037*** (0.0402)	0.1155** (0.0449)						
CONTROL OF CORRUPTION (HIGH)							0.0392 (0.0548)	0.0338 (0.0531)	0.0145 (0.0483)	0.0316 (0.0479)	0.0166 (0.0503)	0.0232 (0.0596)
Adj. R-squared N	0.798 38438	0.798 38438	0.775 20441	0.775 20441	0.786 12869	0.786 12869	0.786 38438	0.786 38438	0.760 20441	0.760 20441	0.771	0.771 12869
Ethnicity fixed effects Pixel area and pop. dens.	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes

Table F.2. RE-ESTIMATING MP14 TABLE VI. This table re-estimates MP14 Table VI, which estimates border regression discontinuity models.

	(1)	(2)	(3)	(4)	(5)	(9)	(7)	(8)	(6)	(10)	(11)	(12)
			Pan	el A: Dep.	variable is	NL						
rule of law (high)	0.0159 (0.0186)	0.0116 (0.0199)	-0.0027 (0.0212)	0.0134 (0.0311)	0.0049 (0.0376)	-0.0207 (0.0330)						
CONTROL OF CORRUPTION (HIGH)							0.0090	0.0133	-0.0030	0.0410	0.0367	0.0024
Adj. R-squared	0.398 19349	0.400 1 03/10	0.400 19349	0.458	0.461	0.462	0.399	0.399	0.400	0.457	0.460	0.461
			Panel B: De	p. variable	e is consum	ption, RF-2)))	
RULE OF LAW (HIGH)	0.1941** (0.0979)	0.1440*** (0.0557)	0.1287** (0.0513)	0.2457 (.)	0.2629*** (0.0890)	0.2007*** (0.0729)						
CONTROL OF CORRUPTION (HIGH)				2			0.2506** (0.0994)	0.1622** (0.0652)	0.1449** (0.0634)	0.3010*** (0.1127)	0.2185*** (0.0743)	0.1869*** (0.0694)
Adj. R-squared N	0.726 17796	0.728 17796	0.728 17796	0.739 8478	0.745 8478	0.748 8478	0.779 18261	0.782 18261	0.783	0.796 9241	0.803	0.804 9241
Ethnicity fixed effects Pixel area and pop. dens.	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes

Table F.3. RE-ESTIMATING MP14 TABLE VII. This table re-estimates MP14 Table VII, which estimates border regression discontinuity models when institutional differences across borders are large.

Additional MP13 models. Tables F.4 to F.7 present results from models in MP13 to demonstrate there is no robust relationship between JURISDICTIONAL HIERARCHY when SED consumption is used as the outcome variable. Each table describes the model being reestimated and the reader can find details about model specifics in MP13.

	(1)	(2)	(3)	(4)	(5)
	Pane	l A: Dep. v	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

Table F4. RE-ESTIMATING MP13, TABLE V PANEL A. This table re-estimates MP13 Table V, Panel A using the MP13 measure oflog of nightlights as outcome variable.

	(1)	(2)	(3)	(4)	(5)
		Panel A:	Dep. varia	able is NL	
Petty Chiefdoms	0.0120	0.0459	0.0290	0.0194	0.0113
	(0.0231)	(0.0346)	(0.0225)	(0.0182)	(0.0137)
Paramount Chiefdoms	0.0538	0.0843*	0.0602*	0.0642**	0.0464**
	(0.0340)	(0.0507)	(0.0311)	(0.0306)	(0.0187)
Pre-Colonial States	0.0853	0.0990*	0.0638**	0.0625***	0.0383**
	(0.0647)	(0.0511)	(0.0252)	(0.0223)	(0.0176)
Adj. R-squared	0.008	0.182	0.268	0.288	0.294
N	61359	61359	61359	61015	61015
	Panel B:	Dep. varia	able is con	sumption f	rom RF-2
Petty Chiefdoms	-0.1631*	0.0487	0.0333	0.0012	0.0070
	(0.0949)	(0.0463)	(0.0439)	(0.0397)	(0.0427)
Paramount Chiefdoms	-0.1040	-0.0147	-0.0368	-0.0221	-0.0096
	(0.1530)	(0.0691)	(0.0626)	(0.0337)	(0.0331)
Pre-Colonial States	-0.0434	0.0076	-0.0246	-0.0425	-0.0505
	(0.1842)	(0.0629)	(0.0696)	(0.0567)	(0.0624)
Adj. R-squared	0.007	0.778	0.793	0.832	0.836
Ν	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
Observations	61359	61359	61359	61015	61015

 Table F.5.
 RE-ESTIMATING MP13, TABLE V PANEL B. This table re-estimates MP13 Table V, Panel B using the transformed consumption variable from RF-2.

Should we get rid of the NL rows in the tables that follow?

	(1)	(2)	(3)	(4)	(5)	(9)	(2)	(8)	(6)
		Panel A:	Dep. varia	ble is NL					
JURISDICTIONAL HIERARCHY	0.0124	0.0077	0.0043	0.0118	0.0084	0.0044	0.0161	0.0115	0.0046
	(0.0094)	(0.0050)	(0.0045)	(0.0115)	(0.0060)	(0.0053)	(0.0168)	(0.0085)	(0.0051)
Adj. R-squared	0.259	0.314	0.323	0.218	0.287	0.295	0.278	0.332	0.343
N	72545	72545	72297	30008	30008	29915	12851	12851	12812
	Panel B:]	Jep. varia	ble is cons	sumption f	from RF-2				
JURISDICTIONAL HIERARCHY	-0.0155	-0.0202	-0.0188	-0.0215	-0.0242	-0.0261	-0.0024	-0.0055	0.0061
	(0.0224)	(0.0247)	(0.0225)	(0.0277)	(0.0305)	(0.0275)	(0.0255)	(0.0283)	(0.0195)
Adj. R-squared	0.855	0.867	0.880	0.869	0.876	0.894	$0.904 \\ 12851$	0.908	0.916
N	72545	72545	72297	30008	30008	29915		12851	12812
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Controls at the Pixel level	No	No	Yes	No	No	Yes	No	No	Yes

Table F.6. RE-ESTIMATING MP13, TABLE VII. The dependent variable is consumption from RF-2.

		Dependent	Variable: Consumption			
	< 100 km of	< 150 km of	< 200 km of			
	ethnic border	ethnic border	ethnic border			
	(1)	(2)	(3)			
Panel A: Pre-Colonial Eth	nic Institutions	and Regional D	evelopment Within Contiguous Ethnic			
Homelands in the Same C	ountry Pixel-Le	vel Analysis in	Areas Close to the Ethnic Border			
Panel 1: Border Th	ickness: Total 50) km (25 km from	each side of the ethnic boundary)			
JURISDICTIONAL HIERARCHY	0.0168	0.0094	0.0105			
	(0.0173)	(0.0158)	(0.0158)			
Adj. R-squared	0.908	0.898	0.897			
Ν	6237	9476	11920			
Panel 2: Border Thi	ckness: Total 10	0 km (50 km fron	n each side of the ethnic boundary)			
JURISDICTIONAL HIERARCHY	0.0117	0.0032	0.0066			
	(0.0178)	(0.0161)	(0.0166)			
Adj. R-squared	0.906	0.896	0.896			
Ν	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 Thi	ckness—Total 50	0 km (25 km fron	n each side of the ethnic boundary)			
JURISDICTIONAL HIERARCHY	0.0104	0.0145	0.0205			
	(0.0272)	(0.0248)	(0.0231)			
Adj. R-squared	0.909	0.899	0.898			
N	6237	9476	11920			
Panel 2: Border Thickness—Total 100 km (50 km from each side of the ethnic boundary)						
JURISDICTIONAL HIERARCHY	0.0128	0.0174	0.0220			
	(0.0397)	(0.0337)	(0.0302)			
Adj. R-squared	0.906	0.896	0.896			
N	4053	7292	9736			
RD-Type Polynomial	Yes	Yes	Yes			
Adjacent-Ethnic-Groups	Yes	Yes	Yes			
Fixed Effects						
Population Density	Yes	Yes	Yes			
Controls at the Pixel level	Yes	Yes	Yes			

Table F.7. RE-ESTIMATING MP13, TABLE VIII USING CONSUMPTION. This table re-estimates the first three models (which use allobservations) from MP13 Table VIII. The dependent variable is consumption from RF-2.

G. References

Battistin, Erich, Blundell, Richard, and Lewbel, Arthur (2009), "Why Is Consumption More Log Normal than Income? Gibrat's Law Revisited," *Journal of Political Economy*, 117, 1140–1154.

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.

Chi, G, Fang, S, and Blumenstock, JE (2022). "Microestimates of Wealth and Poverty for all Low- and Middle-Income Countries," *Proceedings of the National Academy of Sciences*, 119(3), 1-11.

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.

Gibson, J., O. Susan, and G.Boe-Gibson (2020), "Night Lights in Economics: Sources and Uses," *Journal of Economic Surveys* 34, 955–980.

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 wellbeing in Africa," *Nature Communications*, 11, 2583.