Supporting Information

Pokhriyal and Jacques 10.1073/pnas.1700319114

Global MPI

Poverty has traditionally been measured in one dimension, usually income (or consumption)—also known as income poverty. Another internationally comparable measure is the Global MPI, which complements income poverty and is created from nationally representative Demographic and Health Surveys and Multiple Indicator Cluster Survey (DHS-MICS) (42). It was developed by OPHI and the United Nations Development Program. It is a composite of 10 indicators across three critical dimensions—education (years of schooling, school enrollment), health (malnutrition, child mortality), and living conditions (cooking fuel, sanitation, access to drinking water, electricity, floor, asset ownership).

MPI is defined as the percentage of people who are MPI poor (H, headcount of poverty) multiplied by the average intensity of MPI poverty across the poor (A, intensity of poverty). The MPI data for Senegal, used in this study, were downloaded from www.ophi.org.uk/wp-content/uploads/Senegal-2013.pdf.

MPI is robust to decomposition within relevant subgroups of populations, like urban vs. rural, geographic regions (districts/provinces/states), and gender, so that targeted policies can be planned for specific demographics. Countries can also adapt the multidimensional poverty approach to select different indicators and/or update weights that align better with their nation's poverty measure. Countries, like Mexico, Colombia, and Chile, have implemented their own version of national MPI using additional dimensions than MPI, such as employment and social protection, when data are available (www.mppn.org/).

The MPI has some limitations. Though it has been defined from available variables in global surveys (DHS-MICS), some of the potential dimensions of poverty (like gender, income, employment) are not directly incorporated. However, due to the wide availability of these surveys, Global MPI can easily be estimated in more than 100 countries covering 5.2 billion people (42). Consequently, it represents a benchmark index, more interesting than the single dimension poverty line, for replication of this study in another country.

Interpretation of Weights—Along the Dimensions of Poverty

This section complements *Dimensions of Poverty—Interpretation of Weights* and refers to Fig. S2. These interpretations are given for information purposes and are by no means indicators of causality.

Several features related to the presence of water in the commune (water bodies, water, mudflat soil, hydro-morphic soil, and elevation) are positively correlated with water deprivation, although the opposite is observed for the other dimensions. One interpretation is that natural nonportable water would be used for drinking in these areas. On the other hand, access to water is needed for irrigated agriculture, watering livestock, or fishing, all of which can increase income and life quality, which explains the negative relationship for the other dimensions. Interestingly, the distance to a water tower is not quite correlated with this deprivation. Alternatively, the proximity to a water forage would have probably been a more interesting feature.

The food security (access) features (like distance to main roads and urban centers) are also prominent, stressing their importance for development. Millet price has a mixed behavior. Depending on the dimensions, its coefficient is sometimes positive and sometimes negative, without any evident explanation.

The effect of temperature is clear. The higher the maximum temperature and the range, the higher the poverty. Temperature plays a role in crop growth, but it also impacts the environment quality of the people who live in warm (and cold during the night) areas.

The effect of precipitation is less obvious. The amount and the period of rainfall affects the availability of water, which is the main limiting factor in Sahel for crop and forage production. The precipitation seasonality, described by the period during which the water is available, and the precipitation of the warmest quarter (critical period) are logically negatively correlated with poverty. However, the annual precipitation and the precipitation of the wettest month and quarter have a positive coefficient (except for education deprivation). In other words, the more it rains in an area, the poorer it is. The intuition would have been that it was the opposite. But looking more closely, it appears that several features related to agriculture (groundnut production, cassava production, rain-fed croplands) show the same patterns. We interpret that these features define a suitable environment for agricultural areas, which itself is linked to the presence of a rural community tending more to poverty than the urban population.

Understanding Model Uncertainty

The predictive variance associated with the GP model, as calculated using Eq. **4**, indicates the model uncertainty for a test target. The variance does not depend on the observed target values, only on the inputs. The variance at a given test commune is directly related to how many similar communes (in terms of the CDR, environmental, and spatial features) are available in the training data. For instance, if the predictive variance is high for a given test commune, it would mean that the relative density of the training feature vectors in proximity of the feature vector corresponding to the test commune is low, and hence the GP model will yield a higher predictive variance. This could explain the higher variance for MPI predictions observed for rural communes compared with urban communes.

GP Regression Model

The following model is assumed to predict poverty for a commune from a single data source (CDR or environment):

$$y_i = \beta^\top \mathbf{x}_i + f(\mathbf{x}_i) + \epsilon \qquad \textbf{[S1]}$$

where y_i is the target poverty value and $\mathbf{x}_i$ is a vector of independent variables derived from the particular view for the i^{th} commune. Instead of assuming a fixed parametric form for $f()$, we adopt a nonparametric approach, by assuming a GP prior on $f()$, with zero mean function, and kernel function $k()$. The generative process thus becomes:

$$f(\mathbf{x}) \sim GP(0, k(\mathbf{x}, \mathbf{x}'))$$
$$y_i \sim \mathcal{N}(\boldsymbol{\beta}^\top \mathbf{x}_i + f(\mathbf{x}_i), \sigma_n^2), \forall i$$

A GP is a stochastic process, such that any finite sample generated from this stochastic process is jointly multivariate normal (15).

The posterior distribution of $f(\mathbf{x}_*)$ at a test input, $\mathbf{x}_*$, can be computed given a training set of examples, $\{\mathbf{x}_i, f(\mathbf{x}_i)\}_{i=1}^N$.

The joint distribution of the training outputs, $f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots$, and the test output, $f(\mathbf{x}_*)$, according to the GP prior is:

$$\begin{bmatrix} f(\mathbf{x}_1) \\ f(\mathbf{x}_2) \\ \vdots \\ f(\mathbf{x}_N) \\ f(\mathbf{x}_*) \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x_1}) & \ldots & k(\mathbf{x}_1, \mathbf{x_N}) & k(\mathbf{x}_1, \mathbf{x}_*) \\ k(\mathbf{x}_2, \mathbf{x_1}) & \ldots & k(\mathbf{x}_2, \mathbf{x_N}) & k(\mathbf{x}_2, \mathbf{x}_*) \\ \vdots & \ddots & \vdots & \vdots \\ k(\mathbf{x}_N, \mathbf{x_1}) & \ldots & k(\mathbf{x}_N, \mathbf{x_N}) & k(\mathbf{x}_N, \mathbf{x}_*) \\ k(\mathbf{x}_*, \mathbf{x_1}) & \ldots & k(\mathbf{x}_*, \mathbf{x_N}) & k(\mathbf{x}_*, \mathbf{x}_*) \end{bmatrix}\right)$$

For notational simplicity, let K denote a $N \times N$ matrix that contains the kernel computation on each pair of training inputs—that is, $K[i,j] = k(\mathbf{x}_i, \mathbf{x}_j)$—$\mathbf{k}$ be a vector of the kernel computation between each training input and the test input—that is, $\mathbf{k}[i] = k(\mathbf{x}_i, \mathbf{x}_*)$—and k_* be the self-covariance for $\mathbf{x}$—that is, $k_* = k(\mathbf{x}_*, \mathbf{x}_*)$. Moreover, let $\mathbf{f}$ be a $N \times 1$ vector, such that $\mathbf{f}[i] = f(\mathbf{x}_i)$. The above equation can be written as:

$$\begin{bmatrix} \mathbf{f} \\ f(\mathbf{x}_*) \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & \mathbf{k} \\ \mathbf{k}^\top & k_* \end{bmatrix}\right)$$

Since $\mathbf{f}$ and $f(\mathbf{x}_*)$ are jointly Gaussian, one can make use of the well-known Gaussian identity (43) for the conditional distribution of $f(\mathbf{x}_*)$—that is:

$$f(\mathbf{x}_*)|\mathbf{f} \sim \mathcal{N}\left(\mathbf{k}^\top K^{-1}\mathbf{f}, k_* - \mathbf{k}^\top K^{-1}\mathbf{k}\right) \qquad \textbf{[S2]}$$

We assume that the observed poverty for the i^{th} commune, y_i, is equal to the sum of the linear term, the latent function value, with zero mean GP prior, and an independent and identically distributed Gaussian noise ($\sim \mathcal{N}(0, \sigma_n^2)$). Thus, the prior on the observed data will be:

$$\mathbb{E}[y_i] = \beta^\top \mathbf{x}_i$$
$$\text{cov}[y_i, y_j] = k(\mathbf{x}_i, \mathbf{x}_j) + \delta_{ij}\sigma_n^2$$

where δ_{ij} is the Kronecker delta, such that $\delta_{ij} = 1$, if $(i = j)$, and 0 otherwise. For the entire training dataset:

$$\mathbb{E}[\mathbf{y}] = \mathbf{b}$$
$$\text{cov}[\mathbf{y}] = K + \sigma_n^2 I$$

where $\mathbf{b}$ is a N length vector, such that $\mathbf{b}[i] = \beta^\top \mathbf{x}_i$, and I is the $N \times N$ identity matrix. The joint distribution of $\mathbf{y}$ and $f(\mathbf{x}_*)$ can be written as:

$$\begin{bmatrix} \mathbf{y} \\ f(\mathbf{x}_*) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{b} \\ \beta^\top \mathbf{x}_* \end{bmatrix}, \begin{bmatrix} K + \sigma_N^2 I & k \\ k^\top & k_* \end{bmatrix}\right)$$

Using the conditional Gaussian result, similar to Eq. **2**, and noting the relation between y_* and $f(\mathbf{x}_*)$ from Eq. **1**, the conditional distribution for the prediction y_* becomes:

$$\mathbb{E}[y_*] = \beta^\top \mathbf{x}_* + \mathbf{k}^\top (K + \sigma_n^2 I)^{-1}(\mathbf{y} - \mathbf{b})$$
$$\text{var}[y_*] = k_* - \mathbf{k}^\top (K + \sigma_n^2 I)^{-1}\mathbf{k} + \sigma_n^2$$

Estimating Moments of a Mixture Distribution

Let random variable y represent a mixture of two unimodal normal distributions, $y_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $y_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ and mixing probabilities w_1 and w_2, such that $w_1 + w_2 = 1$—that is:

$$y = w_1 y_1 + w_2 y_2$$

Any moment of y can be computed as (44):

$$\mathbb{E}[y^k] = w_1 \mathbb{E}[y_1^k] + w_2 \mathbb{E}[y_2^k]$$

which directly gives:

$$\mathbb{E}[y] = w_1\mu_1 + w_2\mu_2$$

The expression for the variance of y can be derived as follows:

$$\begin{aligned} \text{var}[y] &= \mathbb{E}[y^2] - (\mathbb{E}[y])^2 \\ &= w_1\mathbb{E}[y_1^2] + w_2\mathbb{E}[y_2^2] - (w_1\mu_1 + w_2\mu_2)^2 \\ &= w_1(\text{var}[y_1] + \mu_1^2) + w_2(\text{var}[y_1] + \mu_1^2) - (w_1\mu_1 + w_2\mu_2)^2 \\ &= w_1\sigma_1^2 + w_2\sigma_2^2 + w_1\mu_1^2 + w_2\mu_2^2 \\ &\quad - w_1^2\mu_1^2 - w_2^2\mu_2^2 - 2w_1w_2\mu_1\mu_2 \\ &= w_1\sigma_1^2 + w_2\sigma_2^2 + w_1w_2\mu_1^2 + w_1w_2\mu_1^2 - 2w_1w_2\mu_1\mu_2 \\ &= w_1\sigma_1^2 + w_2\sigma_2^2 + w_1w_2(\mu_1 - \mu_2)^2 \end{aligned}$$

The last result makes use of the fact that $w_1 + w_2 = 1$.

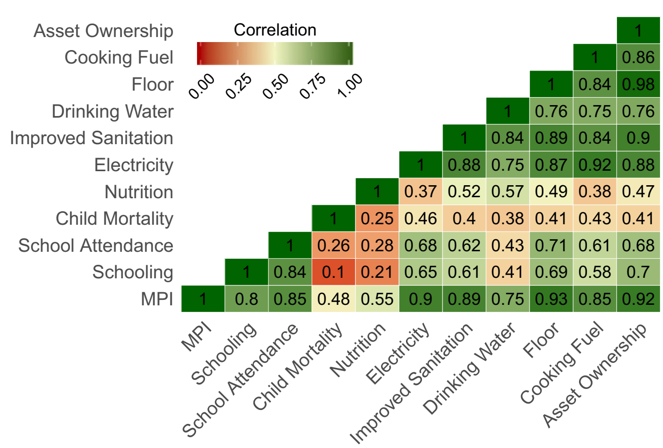

Fig. S1. Spearman correlation matrix between individual deprivations, H (headcount of poverty), A (intensity of poverty), and MPI at the commune level.

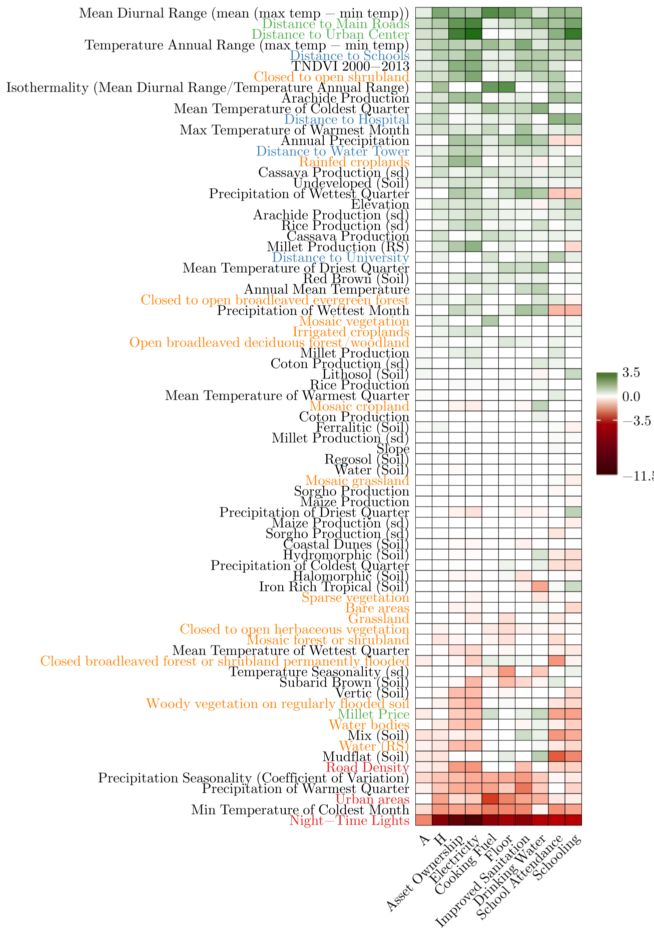

Fig. S2. Visualization of selected features using elastic net regularization on environmental data for prediction of selected deprivations. The rows represent the features, which are ranked according to their weights from positive (marked green) to negative (marked red). Different features groups are color-coded. Features related to food availability are given in black color, whereas those related to food accessibility are colored green. The land cover features are colored yellow, and the features detailing economic activity are in red color. Finally, features depicting access to services are shown in blue. The cells in white were given 0 weights by our model.

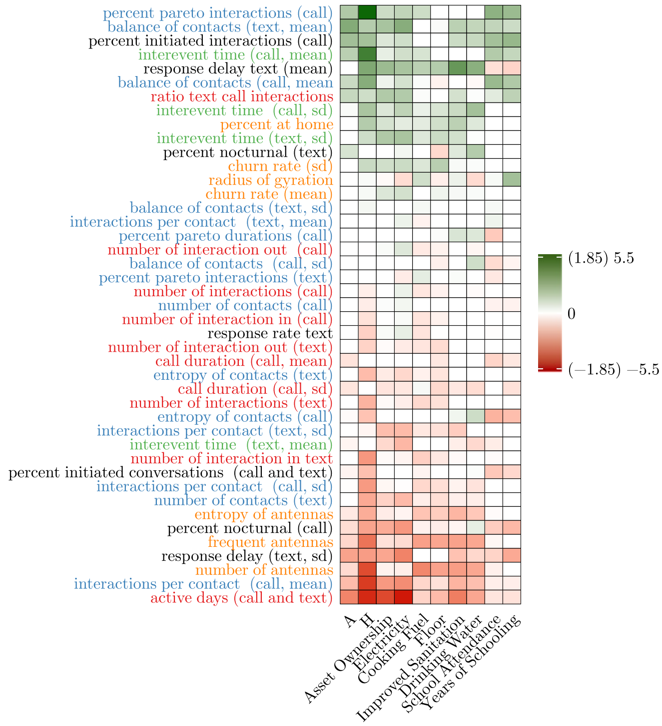

Fig. S3. Visualization of selected features using elastic net regularization on CDR data for prediction of selected deprivations. The rows represent features, which are ranked according to their weights from positive (marked green) to negative (marked red). The columns are the various deprivations. The feature groups are color-coded. Features related to diversity features are colored blue. Those related to spatial aspects are colored yellow. The features related to active behavior are marked in black. The features related to basic phone use are in red, and those related to regularity are in green. The cells in white were given 0 weights by our model. Legend in parentheses corresponds to the different variation in weights. H and A weights vary between 1.85 and −1.85, and for others the weights vary between 5.5 and −5.5.

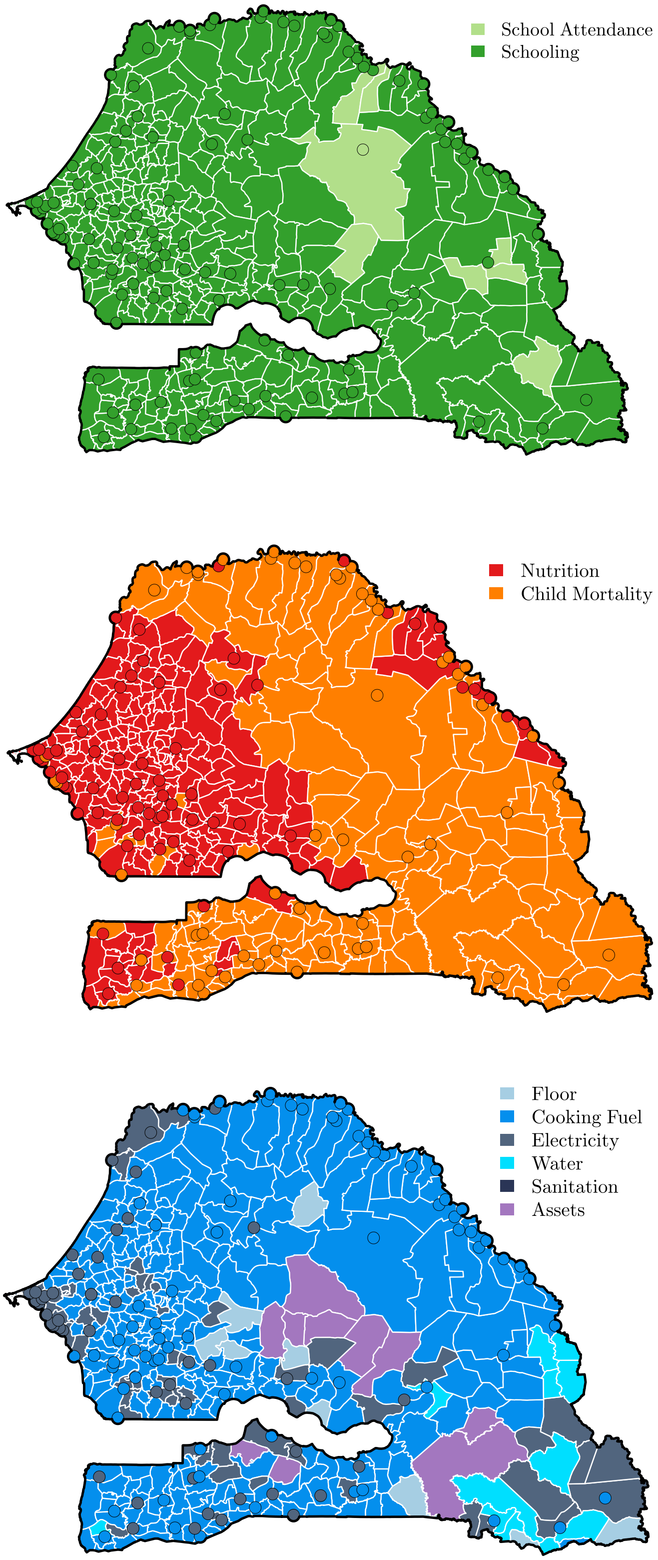

Fig. S4. The highest deprivation by commune as predicted by our model for each dimension of global MPI (from top to bottom: education, health, and standard of living).

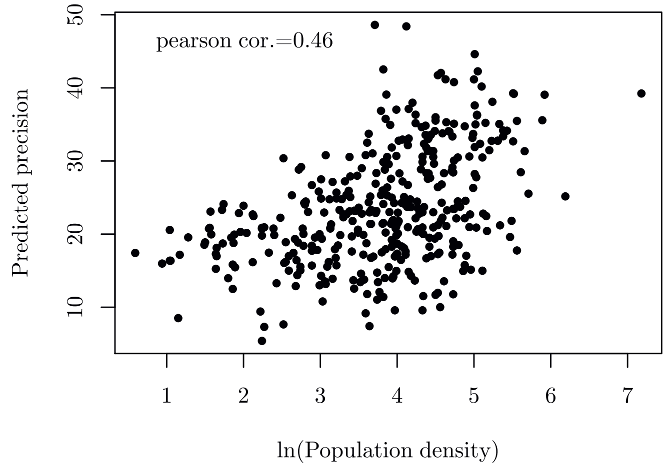

Fig. S5. Relationship between precision of estimates of poverty and the population density of each commune.

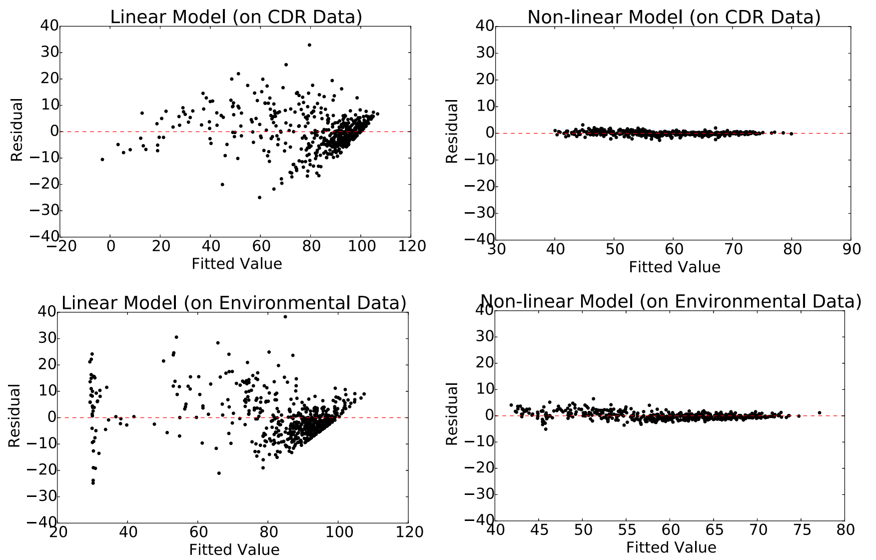

Fig. S6. Residual vs. fit plots to predict incidence of poverty (H) using CDR (*Top*) and environmental (*Bottom*) data. (*Left*) Linear (elastic net regression). (*Right*) Nonlinear (GPR). Linear model fits indicate nonlinearity in the data. The residuals for GPR are normally distributed. Shapiro–Wilk test statistic: CDR, 0.97 (P value $< 10^{-9}$); environmental, 0.95 (P-value $< 10^{-9}$).

Table S1. Source, unit, and expected relationship to poverty of each environmental variable used in this study

Feature (no. of statistics)	Unit	Type of data	Endogeneity	Data sources	Expected relationship to poverty
Food security, availability					
Temperature—annual, annual range, diurnal range, warmest month, warmest quarter, coldest month, coldest quarter, wettest quarter, driest quarter, isothermality (11)	Degree Celsius	Ground	Exogenous	WorldClim database, 1960–1990 (45)	High temperature (+)
Precipitation—annual, wettest month, wettest quarter, driest month, driest quarter, warmest quarter, coldest quarter, coefficient of variation (8)	Millimeter	Ground	Exogenous	WorldClim database, 1960–1990 (45)	Low precipitation (+)
Elevation (1)	Meter	Remote sensing	Exogenous	CGIAR-SRTM data aggregated to 30 s (www.diva-gis.org/)	High elevation (+)
Slope (1)	Degree	Remote sensing	Exogenous	CGIAR-SRTM data aggregated to 30 s	High slope (+)
Soil type (14)	% of territory	Ground	Exogenous	Soil and Terrain Database for Senegal and the Gambia (version 1.0), scale 1:1 million (SOTER Senegal Gambia, www.isric.eu/projects/soter-senegal-and-gambia)	Poor agronomic soil (+)
NDVI (2)	—	Remote sensing	Endogenous	10-d temporal synthesis of 1 km SPOT-VEGETATION satellite images (2000–2013) (www.vgt.vito.be)	Low NDVI (in rural areas) (+)
Crop production (7)	Ton	Ground	Endogenous	Direction de Analyse, de la Prévision et des Statistiques Agricoles (DAPSA) 2000–2014 database (46)	Low production (in rural areas) (+)
Food security (access)					
Millet price (1)	CFA franc/kilogram	Ground	Endogenous	Modeling based on local supply and demand (47)	High millet price (+)
Proximity to urban centers, Market (1)	Kilometer	GIS	Endogenous	ANSD	Far from urban centers (+)
Proximity to main roads (1)	Kilometer	GIS	Endogenous	Open Street Map (www.openstreetmap.org)	Far from main road (+)
Economic activity					
Nighttime lights (2)		Remote sensing	Endogenous	Version 4 of the 2013 nighttime lights time series captured by the Operational Linescan System of the Defense Meteorological Satellite Program (stable lights)	Low density of of light (+)
Density of roads (1)	Kilometer	GIS	Endogenous	Open Street Map	Low density of roads (+)
Land cover					
Land cover (20)	% of territory	Remote sensing	Exogenous/ endogenous	2005 1:100,000 scale Senegal Land Cover Map produced by the Global Land Cover Network (48) based on GlobCover 2005 map (33)	Urban areas (−), cropland (+), forest (+), grassland (+)
Access to facilities					
Proximity to school/university (1)	Kilometer	GIS	Endogenous	Open Street Map	Far from school/ university (+)

Table S1. Cont.

Feature (no. of statistics)	Unit	Type of data	Endogeneity	Data sources	Expected relationship to poverty
Proximity to water tower (1)	Kilometer	GIS	Endogenous	Open Street Map	Far from water tower (+)
Proximity to hospital (1)	Kilometer	GIS	Endogenous	Open Street Map	Far from hospital (+)
Total	81				

Table S2. List of core features extracted for each individual from CDR data using the Bandicoot toolbox (31)

Features (no. of statistics)	Description
Regularity	
Interevent time (4)	The interevent time between two records of the user.
Diversity	
Number of contacts (2)	The number of contacts with whom the user interacted (call and text handled separately).
Entropy of contacts (2)	The entropy of the user's contacts, both for call and text.
Balance of contacts (4)	The balance of interactions per contact. This feature is calculated—each for text and call. For every contact, the balance is the number of outgoing interactions divided by the total number of interactions (in + out).
Interactions per contact (4)	The number of interactions a user had with each of his or her contacts.
Percent pareto interactions (2)	The percentage of user's contacts that account for 80% of his or her interactions.
Percent pareto durations (1)	The percentage of user's contacts that account for 80% of his or her total time spend on the phone.
Active behavior	
Percent nocturnal (2)	The percentage of interactions the user had at night (call and text).
Percent initiated conversations (1)	The percentage of conversations that have been initiated by the user both for call and text.
Percent initiated interactions (1)	The percentage of calls initiated by the user.
Response delay (2)	The response delay of the user within a conversation (in seconds). This is calculated for text (SD and mean of the response delay).
Response rate (1)	The response rate of the user (between 0 and 1).
Basic phone use	
Active days (1)	The number of days during which the user was active.
Call duration (2)	The SD and the mean of the duration of user's calls.
Number of interactions (6)	The number of interactions.
Ratio of text and call interactions (1)	This computes the ratio of the text and call interactions.
Spatial behavior	
Number of antennae (1)	The number of unique places visited.
Entropy of antennas (1)	The entropy of visited antennas.
Percent at home (1)	The percentage of interactions the user had while he or she was at home.
Radius of gyration (1)	Returns the radius of gyration, the equivalent distance of the mass from the center of gravity, for all visited places.
Frequent antennas (1)	The number of locations that accounts for 80% of the locations where the user was.
Churn rate (2)	The SD and mean of the frequency spent at every antenna each week.
Total	43

Features are grouped into categories based on prior research (29). These features are calculated for each month, so in total there are $43 \times 12 = 516$ features.

Table S3. Brief review of poverty estimation methods based on environmental data

Ref.	Poverty variable	Model	Important variables	Main conclusions	Region
(49)	Daily consumption expenditure	Regression, correlation	Indoor air pollution (wood/charcoal use), access to clean water, no sanitation, diarrhea, outdoor air pollution (number of deaths from PM10)	Substantial variability across countries	Cambodia, Lao PDR, Vietnam
(50)	Per capita income	Regression	Mean road density, share in in internal revenue allotment, agrarian reform accomplishment rate, population growth, distance to major cities, mean elevation, percentage of slope with agricultural limitations, and mean annual rainfall	Spatial variation in poverty is mainly caused by disparities on access to road infrastructure	Philippines
(51)	Food expenditure	Regression and clustering	Proportion of irrigation land, average landholding sizes	Poverty maps show significant spatial clustering of poor and nonpoor areas	Sri Lanka
(11)	Per capita expenditure	Spatial regression	Slope, soil type, distance/ travel time to public resources, elevation, type of land use, demographic variables	Increasing access to roads and improving soil conditions would result in decline in poverty	Kenya
(52)	Per capita expenditure	Regression	Distance to town, soil quality, slope	Poverty in the remote areas is linked to low agricultural potential and lack of market access	Vietnam
(53)	Household expenditure	Discriminant analysis	Distance to market, agro-climatic variables, diseases risk, livestock density	Satellite-derived variables tended to dominate the list of selected variables that determine poverty predictions	Uganda
(8)	Household consumption expenditure, asset wealth	Transfer learning (deep learning)	Roofing material, distance to urban areas	Interesting potential of machine learning method using limited training data	Nigeria, Tanzania, Uganda, Malawi, Rwanda
(54)	Household expenditure	Spatial regression, geographically weighted regression	Crop diversity, education, nonagricultural economic activities	Spatial nonstationarity of the relationship between poverty and its determinants	Malawi
(55)	Household income	Geographically weighted regression	Education, accessibility, and services	High poverty incidence that corresponds with ecologically depressed areas. However, other livelihood-influencing factors such as education, accessibility, and services are significantly correlated with poverty.	Bangladesh
(56)	Relative welfare (female literacy, land ownership, deprived class, and water source)	Random forests	Travel time to market towns, percentageof a village covered with woodland, and percentage of a village covered with winter crop	Satellite sensor data are strongly associated with aspects of rural welfare for an extensive region of a developing country	India

Table S4. Brief review of poverty estimation methods based on CDR data

Ref.	Data source	Model (number of features)	Sample size	Time period	Results, Pearson's r	Spatial resolution of validation	Poverty measure	Region
(7)	CDR and phone survey	Linear regression (5,088)	1.5 M (CDR) + 856 (survey)	9 mo	0.68	492 DHS clusters	DHS composite wealth index	Rwanda
(23)	CDR	Support vector machine (279)	500 K	6 mo	0.80	—	Socioeconomic levels	Urban area in a Latin American city
(57)	CDR	Linear regression (OLS) (5)	5 M and 928 K	20 wk and 6 wk	—	11 subprefecture level (CIV)	IMF poverty rate	Cote d'Ivoire and anonymous region B
(58)	CDR	Linear regression (33)	9 M and 150 K	12 mo	0.82	14 regions in Senegal	MPI (OPHI)	Senegal

Table S5. Spatially-cross validated results of the predictions of MPI, incidence of poverty (H), and intensity of poverty (A), along with the individual indicators for poverty given by our model using disparate datasets

	Multisource data			CDR			Environment			Concatenated		
Poverty indicator	Corr.	Rank corr.	RMSE	Corr.	Rank corr.	RMSE	Corr.	Rank corr.	RMSE	Corr.	Rank corr.	RMSE
MPI	0.91 (0.06)	0.88 (0.06)	0.08 (0.01)	0.89 (0.07)	0.86 (0.07)	0.08 (0.01)	0.84 (0.09)	0.80 (0.10)	0.10 (0.02)	0.90 (0.06)	0.85 (0.07)	0.10 (0.02)
H	0.91 (0.07)	0.85 (0.08)	10.79 (3.96)	0.90 (0.08)	0.84 (0.08)	10.76 (2.60)	0.83 (0.11)	0.75 (0.11)	13.65 (4.86)	0.90 (0.07)	0.83 (0.08)	11.34 (3.87)
A	0.86 (0.05)	0.85 (0.07)	04.71 (0.96)	0.83 (0.07)	0.82 (0.08)	04.98 (1.14)	0.81 (0.07)	0.79 (0.08)	05.36 (0.75)	0.84 (0.07)	0.82 (0.08)	5.52 (1.40)
Individual indicators of poverty												
Education												
Years of schooling	0.85 (0.04)	0.85 (0.04)	12.00 (1.21)	0.81 (0.05)	0.80 (0.06)	13.30 (1.55)	0.76 (0.07)	0.75 (0.08)	15.42 (2.48)	00.85 (0.04)	0.84 (0.04)	12.06 (01.01)
School attendance	0.86 (0.05)	0.83 (0.06)	11.68 (1.83)	0.82 (0.07)	0.81 (0.07)	12.85 (1.73)	0.75 (0.09)	0.72 (0.09)	14.54 (3.06)	0.85 (0.05)	0.83 (0.06)	11.60 (2.05)
Health												
Child mortality	0.45 (0.15)	0.46 (0.16)	10.91 (0.58)	0.45 (0.13)	0.48 (0.13)	11.32 (00.73)	0.34 (0.19)	0.33 (0.21)	11.54 (0.65)	0.45 (0.14)	0.45 (0.16)	10.85 (0.49)
Nutrition	0.52 (0.15)	0.53 (0.15)	14.61 (3.65)	0.54 (0.11)	0.55 (0.11)	14.49 (3.10)	0.38 (0.26)	0.37 (0.25)	16.28 (3.99)	0.47 (0.21)	0.46 (0.22)	15.33 (4.24)
Standard of living												
Cooking fuel	0.86 (0.14)	0.70 (0.18)	13.82 (8.76)	0.83 (0.14)	0.68 (0.16)	12.98 (7.00)	0.76 (0.20)	0.58 (0.25)	16.49 (8.78)	0.86 (0.13)	0.70 (0.18)	15.56 (9.19)
Sanitation	0.79 (0.17)	0.70 (0.18)	16.99 (3.42)	0.74 (0.17)	0.69 (0.17)	18.05 (3.14)	0.72 (0.22)	0.61 (0.26)	18.64 (4.33)	0.77 (0.20)	0.66 (0.23)	18.69 (3.91)
Water	0.75 (0.14)	0.72 (0.14)	14.60 (3.22)	0.74 (0.13)	0.71 (0.12)	14.70 (2.98)	0.67 (0.20)	0.61 (0.21)	16.97 (3.25)	0.68 (0.21)	0.62 (0.22)	17.15 (3.20)
Electricity	0.88 (0.04)	0.84 (0.07)	15.09 (0.98)	0.86 (0.04)	0.83 (0.06)	16.67 (1.25)	0.79 (0.10)	0.72 (0.13)	20.27 (1.72)	0.84 (0.05)	0.80 (0.09)	18.61 (1.65)
Floor	0.78 (0.15)	0.68 (0.14)	15.79 (5.79)	0.79 (0.13)	0.70 (0.12)	15.24 (4.93)	0.64 (0.24)	0.54 (0.23)	17.87 (6.22)	0.74 (0.19)	0.63 (0.16)	16.58 (5.81)
Asset ownership	0.89 (0.04)	0.86 (0.05)	12.61 (1.33)	0.87 (0.04)	0.85 (0.04)	13.81 (1.20)	0.80 (0.11)	0.75 (0.11)	17.05 (2.69)	0.85 (0.05)	0.82 (0.06)	15.37 (1.48)

The results are compared with models learned on single source and on concatenated feature space. Corr., Pearson's *r* correlation; rank corr., Spearman's rank correlation; RMSE, root mean square error. For both types of correlations, all *p* values were less than 10^{-20}. An SD associated with the multiple runs for each measurement is reported within parentheses.

Table S6. List of the important features chosen by our model to predict each of H, A, schooling, school attendance, cooking fuel, sanitation, water, electricity, floor, and assets

Feature type	H	A	Schooling	School attendance	Cooking fuel	Sanitation	Water	Electricity	Floor	Assets
Basic										
Active days call and text	–	–			–	–	–	–	–	–
Ratio of call/text interactions			+			+		+		+
Number of interactions in text					–					
Regularity										
Interevent time call, mean	+	+	+	+						
Interevent time call/text, SD						+	+	+	+	+
Diversity										
Balance of contacts text, mean		+	+	+		+	+	+		+
Percent pareto interactions call	+	+	+	+	+			+		+
Interactions per contact call, mean	–	–			–	–	–	–		–
Entropy of contacts call			–	–				–		
Active										
Response delay text, mean	+		–		+	+	+	+	+	+
Response delay text, SD		–	–	–		–	–	–		–
Percent initiated interactions, call		+	+	+		+	+			
Percent initiated conversations, call and text			–	–						
Spatial										
Frequent antennas	–				–	–	–		–	
Number of antennas	–				–	–	–		–	
Radius of gyration			+		+					
Entropy of antennas					–	–	–		–	

The features having positive relationships with the various deprivations are marked as + in the cell corresponding to the feature name and the deprivation. Otherwise they are marked as –. The various semantic groupings under which the different features fall are also listed.

Table S7. A summary of poverty indicators and associated deprivations, with emphasis on how our methodology calculates them using the RGPHAE census data, keeping in view the OPHI guidelines

Poverty indicators	Deprivation standards of a household used by OPHI for MPI calculation	RGPHAE census questionnaire response used by our methodology for MPI calculation
Health		
Child mortality	At least one child has died	About living and deceased children in the household
Nutrition	Any member is undernourished	About going hunger at night for the past few months
Education		
School attendance	Any school-aged child is not attending school up to grade 8	About school-aged currently not in school
Years of schooling	No member who has completed at least 5 y of education	About higher schooling of any member
Standard of living		
Cooking fuel	Uses solid fuels for cooking	Household does not use electricity or natural gas for cooking
Electricity	No access to electricity	No electricity or generator
Sanitation	No access to adequate sanitation or if it is shared	Household has no sewer connection or pit
Drinking water	No access to safe drinking water	No water tap in household
Flooring	Has dirt/earth/dung floor	Household has dirt/earth/dung floor
Assets	Has only one small asset (radio, TV, refrigerator, phone, bicycle, motorbike) and it has no car	Household has one asset (radio, TV, refrigerator, phone, bicycle, motorbike) and it has no car

Table S8. Comparative table showing how our model performs compared with only nightlights and a previous work (used as a baseline) using only four features—namely, call volume and mobile ownership per capita, nightlights, and population density

Data source	Model	Results, Pearson's r
Nightlights	Linear regression	0.39
(20)	Linear regression	0.84
Our model	GP regression	0.91