Nighttime lights as a proxy for human development at the local level*

Anna Bruederle† and Roland Hodler‡

July 4, 2017

Abstract

Nighttime lights are increasingly used by social scientists as a proxy for economic activity and economic development in subnational spatial units. However, so far, our understanding of what nighttime lights capture is limited. We construct local indicators of household wealth, education and health from geo-coded Demographic and Health Surveys (DHS) for 29 African countries. We show that nighttime lights are positively associated with these indicators across DHS cluster locations as well as across grid cells of roughly 50×50 km. We conclude that nighttime lights are a good proxy for human development at the local level. (96 words)

Keywords: Nighttime lights, local development, Africa.

JEL classification: I15, I25, I32, O15, O55.

*We acknowledge financial support by the Basic Research Fund of the University of St.Gallen.
†University of St.Gallen. Email: mirjamanna.bruederle@unisg.ch.
‡University of St.Gallen, CEPR, CESifo, and OxCarre. Email: roland.hodler@unisg.ch.

1 Introduction

Economic and social data for subnational administrative regions, such as provinces, districts or municipalities, are unavailable for most developing countries, and of poor quality if they exist. For geographical units that cut across administrative borders, such as ethnic homelands or ecological zones, reliable economic and social indicators are even rarer. This lack of spatially disaggregated data has inhibited empirical research on important questions in economics and related social sciences for many years. In the absence of (reliable) subnational data, social scientists have recently resorted to alternative measures that do not depend on data collection on the ground. The most prominent example are nighttime lights. Data on nighttime light intensity are calculated from weather satellite recordings and available from the U.S. National Oceanic and Atmospheric Administration (NOAA) as an annual time series with global coverage. A key benefit of these data is the high spatial resolution with pixels corresponding to less than one square kilometer, which allows researchers to aggregate these data at the level of the subnational units they want to study. In addition, nighttime lights are measured with consistent quality across countries with very different institutional capacities, and are not susceptible to politically motivated manipulation.

Early studies that proposed nighttime lights as a proxy for economic outcomes include Sutton and Costanza (2002), Doll et al. (2006), Sutton et al. (2007), Noor et al. (2008), Elvidge et al. (2009), and Ghosh et al. (2009). It was however the seminal contribution by Henderson et al. (2012) that led to the broad recognition in economics and related social sciences that nighttime lights can serve as a proxy for economic activity or economic development. Since then, nighttime lights have been used to measure economic activity or economic development within administrative regions (e.g. Hodler and Raschky, 2014a,b), cities and municipalities (e.g. Brown et al., 2016; Storeygard, 2016), ethnic homelands (e.g. Michalopoulos and Papaioannou, 2013, 2014; De Luca et al., 2015; Alesina et al., 2016), and grid cells of various sizes (e.g. Besley and Reynal-Querol, 2014; Montalvo and Reynal-Querol, 2016; Storeygard, 2016; Henderson et al., 2017). Thereby, nighttime lights have helped to improve our understanding of comparative development (Michalopoulos and Papaioannou, 2013, 2014; Besley and Reynal-Querol, 2014; Henderson et al., 2017) as well as a host of other topics ranging from regional favoritism (Hodler and Raschky, 2014b) to ethnic inequality (Alesina et al., 2016) and micro-finance (Brown et al., 2016).[1]

We expect that the use of nighttime lights in the social sciences will continue to grow. Many more social scientists are acquiring the skill set necessary to work with spatial data, including knowledge in geographical information system (GIS) software. In addition, nighttime lights have been added to the PRIO-GRID dataset by Tollefsen et al. (2012), which provides time series of a rich set of variables in a standardized spatial grid structure with

[1] Donaldson and Storeygard (2016) review applications of nighttime lights and other satellite data in economics.

cells of 0.5×0.5 decimal degrees.[2] This facilitates the use of nighttime lights data, especially for researchers without GIS skills.

The surge in the use of nighttime lights contrasts with our rather limited understanding of what nighttime lights do and do not capture. Henderson et al. (2012) show that changes in nighttime lights correlate with growth in GDP at the level of countries. Hodler and Raschky (2014b) demonstrate that this correlation holds also for subnational administrative regions based on the province-level GDP data by Gennaioli et al. (2014). Chen and Nordhaus (2011) document a positive relationship between nighttime lights and GDP at the level of grid cells of 1×1 decimal degrees. These studies however leave at least two important gaps in the evidence that underpins the validity of nighttime light as a tool in the social sciences. First, we do not know whether nighttime lights are an accurate proxy for economic activity and economic development for small spatial units such as municipalities. Second, we do not know whether nighttime lights also capture other important dimensions of human development such as education and health.

We aim to fill these gaps. We explore the relationship between nighttime lights and human development at the local level in Africa. We focus on Africa where the lack of disaggregated data to measure human development is particularly eminent. We use georeferenced data from the Demographic and Health Surveys (DHS) to construct local measures of household wealth, education and health, and relate them to nighttime lights. We use two different spatial units for our analysis, differing in their resolution: circular zones with a radius of up to 5 km around DHS cluster locations, and PRIO-GRID cells.

We find that more intense nighttime lights are associated with better human development outcomes in terms of household wealth, education and health. This association holds both across our small circular zones and across the larger PRIO-GRID cells. Further, these associations hold when we compare spatial units within a country and when we compare them over the entire African continent. We also document that nighttime lights contain information about variation in local human development after controlling for local population density, urbanization and electrification. We conclude that nighttime lights are a good proxy not only for economic development, but also for other dimensions of human development related to education and health, measured within small spatial units such as municipalities or within somewhat larger spatial units such as PRIO-GRID cells.

Our study contributes to the literature on the relation between nighttime lights and purely economic outcomes. The contributions by Chen and Nordhaus (2011), Henderson et al. (2012), and Hodler and Raschky (2014b) all focus on the relation with GDP in relatively large spatial units. The first study to document the correlation between nighttime lights and a survey-based wealth index is Noor et al. (2008), who focuses on provinces. More recently, Weidmann and Schutte (2017) show that nighttime lights are a good predictor of wealth as measured by the DHS wealth index at the local level. Furthermore, Jean

[2]One decimal degree corresponds to approximately 110 km at the equator.

et al. (2016) show that accurate estimates of consumption and wealth at the level of survey clusters can be produced by a machine learning technique that processes daytime satellite images and nighttime lights. The most fine-grained analysis of how nighttime lights are associated with economic outcomes is provided by Mellander et al. (2015) in a study on Sweden. They use geo-coded data on population, enterprises, employment and wages at a resolution of 250×250 meter grid cells in urban and 1×1 km grid cells in rural areas. Data at this level of spatial accuracy are however typically not available for developing countries, which is why nighttime lights are most promising for these countries. None of these studies has looked at the relation between nighttime lights and economic outcomes across PRIO-GRID cells or similarly sized spatial units.

Little is known about the relation between nighttime lights and human development more broadly defined. Chen (2015) finds that nighttime light is correlated with infant mortality and poverty rates at the level of provinces, but we are not aware of any study that looks at various human development outcomes, e.g., related to education, and at a higher spatial resolution. This may be surprising given that the question about the relation between nighttime lights and human development is reminiscent of the old debate on whether GDP per capita is a proxy not only for economic development narrowly defined, but also for human development more broadly defined.[3]

The remainder of the paper is structured as follows: Section 2 discusses the data and how we process them. Section 3 presents the empirical specification, and Section 4 our findings. Section 5 briefly concludes.

2 Data

We use the nighttime lights data by NOAA and combine them with 71 geo-coded Demographic and Health Surveys (DHS) from 29 African countries for the years 1992 to 2013. Before discussing these data in detail, we introduce the spatial units that we use in our analysis. First, we use circular zones around reported center points of DHS clusters, with a 2 km radius for urban clusters and a 5 km radius for rural clusters. Second, we use grid cells of 0.5×0.5 decimal degrees, aligned to the PRIO-GRID (Tollefsen et al., 2012). Our samples comprise around 29,000 circular zones and 7,500 PRIO-GRID cells.

Figure 1 illustrates these two different spatial units for Western Kenya.

Figure 1 around here

The dots represent the reported DHS cluster center points and the (red) circles around these dots our circular zones. The (green) rectangular cells are the PRIO-GRID cells. These spatial units are clipped at country borders (as can be seen on the left in Figure 1)

[3]See Sen et al. (2010) for a recent contribution to this old debate.

and costal lines. Brighter tones of grey in the background indicate more intense nighttime lights.

2.1 Nighttime lights as the main explanatory variable

Our nighttime lights variable is based on satellite images collected by the Operational Linescan System sensors installed on satellites of the Defense Meteorological Satellite Program. These weather satellites circle the earth several times per day and collect a digital stream of images relevant for weather observation and forecasting. The sensors are designed to help identify cloud coverage at night through detecting moonlight reflections, but on cloud-free nights they record light emissions from the earth's surface. These images are processed by NOAA into global annual composites of cloud-free nighttime lights.

Following most of the economic literature, we choose the Stable Lights series from among the NOAA's various nighttime light products for our analysis. This is a series of global maps showing the relative nighttime light intensities on the earth's surface averaged over calendar years, where transient lights that are deemed ephemeral have been filtered out and non-lit areas are set to zero. Baugh et al. (2013) describe the steps of screening and filtering of the raw images, both manual and by automated algorithms, that NOAA implements to obtain the Stable Lights: First, from all the images collected throughout a year, only nighttime observations not affected by lunar illuminance are selected to feed into the annual composite. Further, images affected by cloud coverage are removed because the presence of clouds can either obscure lights completely or diffuse the signals so that they appear larger but dimmer. As a next step, transient light is separated from persistent light in order to filter out lights from, e.g., forest fires or fishing boats. This is done by an algorithm that analyzes the composite histograms and removes bright outliers. The selected segments are then averaged into a global annual cloud-free outlier-removed image. This annual average still contains background noise, i.e., non-zero light intensity values in areas where lights are not present. Therefore, as a final step, areas that are below a locally computed threshold for background noise are deemed as areas where lights are not present and set to zero.

The NOAA currently provides annual data for the time period from 1992–2013, in gridded format with pixels of 30×30 arc seconds. This pixel size corresponds to less than one square kilometer at the equator. For each of the pixels, annual average light intensity is reported in digital numbers (DN) ranging from 0 to 63, with higher values implying more intense nighttime light. We construct from these data our nighttime lights variable (labelled light) as the average DN of all nighttime light pixels within a spatial unit, i.e., a small circular zone or a PRIO-GRID cell, for the year in which the respective DHS survey was carried out.[4]

[4]If DHS data collection extended over two calendar years, we use the earlier year. The reason is that most of the survey questions we use to construct our development indicators refer to the recent past.

The Stable Lights series has two caveats that can be a concern for applications in the social sciences. First, DN are often top-coded at 63 in the centers of metropolitan areas due to calibration of the sensors which allows to detect very low levels of illuminance. The Stable Lights maps hence do not allow to distinguish between bright urban centers and their periphery.[5] However, for the African continent the fraction of pixels in the original data with DN 63 is less than 0.06 percent, which is why we do not consider it a concern for our study and other applications focusing on Africa.

The second caveat is that permanent light sources of low intensity might be inappropriately set to zero by the processing steps involved in the preparation of the Stable Lights series. This concern has been pointed out by Henderson et al. (2012), who note that pixels with DN 1 and 2 seem to be underrepresented in the data. The question arises whether we should interpret a DN of zero as complete absence of light emissions or as possibly very low light emissions that has been filtered away. In our sample, 51 percent of the circular zones around reported DHS cluster center points contain only pixels with a DN of zero. The same holds true for 42 percent of the PRIO-GRID cells. Given that we know that all spatial units in our sample are inhabited, we interpret a DN of zero as very low nighttime light emissions.

2.2 DHS-based development indicators as dependent variables

Our dependent variables are constructed from the Demographic and Health Surveys (DHS), which are large periodic household surveys that have been carried out in low-income countries in Africa and elsewhere since the 1980s (ICF International, 1992-2013). These surveys primarily collect information from women at childbearing age on a wide range of topics related to health, nutrition, fertility and education, as well as a set of household characteristics such as access to infrastructure and ownership of household assets. In each country, households are selected to produce nationally representative samples. Usually, sampling is done through a stratified cluster design, based on the country's most recent population census, in two stages. At the first stage, clusters are drawn from official listings of census enumeration areas, which in most countries correspond to small villages or blocks within larger villages or cities. At the second stage, a sample of households is drawn randomly from a list of households in each cluster. Mean outcomes for a cluster should thus provide an accurate measure of the local level of development.

For those DHS that are geo-referenced, the data contain geo-coordinates of the cluster center points, usually recorded with Global Positioning System (GPS) receivers. The actual location of each household is not recorded, and the DHS provide no indication how far households are scattered around these cluster center points. To further ensure confidentiality of the respondents, some noise is added to the coordinates by displacing each cluster center point in a random direction and by a random distance of 0–2 km for

[5] Letu et al. (2012) provide a detailed discussion of the top-coding problem.

urban clusters, and 0–5 km for rural clusters, with 1 percent of rural clusters displaced by up to 10 km. This displacement is the reason why we choose circular zones with a radius of 2 km for urban clusters and 5 km for rural cluster as the high spatial resolution units in our analysis.

Our sample includes all 71 survey waves with GPS-measured geo-coordinates that have been carried out in African countries over the time period for which nighttime light data are currently available, i.e., 1992–2013.[6] These survey waves were conducted across 29 different countries from all over Africa.[7] DHS follow a largely consistent methodology and structure across countries and years, which is why they qualify for an analysis of local development in a sample with many countries and survey waves.[8]

In order to measure human development in our sample localities, we look at the three key dimensions of human development as reflected in the Human Development Index: a decent standard of living, education, and health. For each of these dimensions, we construct two indicators for which the DHS in our sample provide the relevant data.

We measure standards of living based on household wealth, partly because DHS do not collect data on household income or consumption. As our main wealth measure, we use the DHS wealth index when it is available. The DHS wealth index is constructed as a linear combination of indicators of whether the household owns selected assets; of the type of water, sanitation and energy facilities; and of the housing quality. Weights for each of the components are derived by principal component analysis (PCA).[9] For surveys for which the DHS wealth index is not available, we compute an analogous index following the DHS methodology.[10] Given that the index score is not meaningful per se, households are categorized into five bins separated by quintiles on the distribution of the index score within the survey wave. Depending on the bin into which the household falls, it is assigned a rank order between one and five, where one is assigned to the poorest fifth, and five is assigned to the wealthiest fifth of households within the survey wave. We use these rank orders as indicators of the household's relative wealth.

Several of the components of the DHS wealth index depend on electricity access for their use, and electricity access is inherently correlated with nighttime lights (see Section

[6]We exclude from our analysis DHS clusters whose location has been identified through gazetteers. This is because gazetteers identify locations by the center of the respective village or city rather than by the center of the respective block in case of larger villages or cities. Clusters located by gazetteers make up less than 1 percent of all geo-coded DHS clusters in Africa in our sample period.

[7]Table A.1 in Appendix A provides a list of these 29 countries and indicates the number of survey waves for each country in our sample.

[8]Nevertheless, amendments to the questionnaires are made from time to time, such that the coding of some development indicators needs to be country-specific.

[9]For a detailed description of how the DHS wealth index is constructed, see Rutstein and Johnson (2004). For a general discussion on the use of asset indices to capture household wealth, see, e.g., Filmer and Scott (2012).

[10]We thereby include those wealth indicators that are available across all surveys: ownership of radio, television, refrigerator, bicycle, motorcycle, and car; floor materials; type of drinking water source; type of toilet facility; and access to electricity.

3). We therefore construct a second wealth index (labelled e-free wealth), again following the DHS wealth index methodology, in which we include only indicators that do not depend on electricity.[11] Again, we categorize all households of a survey wave into five equally sized bins and use the bin rank order as indicators of the household's relative e-free wealth. Note that the methodology used to construct these two wealth measures does not allow for comparisons of wealth across countries.

We measure education using primary school attendance and the number of years of schooling. One can interpret the former as a flow and the latter as a stock measure of education. We calculate the net attendance ratio in primary education (labelled school attendance) as the ratio of the number of children of official primary school age (as defined by the national education system) who have attended primary school in the year preceding the survey to the total population of children of official primary school age. The number of years of schooling is based on all household members aged 18 or older.

We measure health using the infant mortality rate (labelled infant mortality) and the proportion of births attended by skilled health personnel (labelled birth assistance). The former is a commonly used measure of health and the latter a measure of access to health services. The infant mortality rate is defined as the number of infants that died before reaching the age of one year per 1,000 live births. In order to make full use of records for children born within a relatively short period preceding the survey, we follow the DHS methodology and derive infant mortality rates through a synthetic cohort life table approach. This approach calculates mortality probabilities for small age segments (0 months, 1-2 months, 3-5 months, 6-12 months) and combines them into the one-year age segment. We use birth records from the three years preceding the survey, because we aim to generate a relatively near-term picture of infant mortality, but require a reasonable number of births to compute the infant mortality rate. Birth assistance is the percentage of deliveries attended by a doctor, a nurse or a midwife among all births in the three years preceding the survey.

For each of these indicators, we first derive shares or mean values for each single DHS cluster, which we use as the value of the dependent variables in our analysis based on circular zones around DHS clusters.[12]

Second, we aggregate the human development indicators at the level of PRIO-GRID cells. We therefore intersect our 2 km or 5 km circular zones around the DHS cluster center points with the PRIO-GRID cells (using GIS). Observations from DHS clusters whose reported center point lies more than 2 km (for urban clusters) or more than 5 km (for rural clusters) from cell boundaries are entirely allocated to their respective cell. For

[11]These electricity-free indicators are: ownership of bicycle, motorcycle, and car; floor materials; type of drinking water source; and type of toilet facility.

[12]The number of observations available at a DHS cluster varies across our indicators depending on the relevant base, i.e., households, children at school age, adults over age 18, or births in the three years preceding the survey.

other DHS clusters we allocate the available observations to the different nearby cells in proportion to the area share of the circular zone that falls within these cells. This procedure yields cell-level values for our development indicators for all grid cells that intersect with at least one of our circular zones. The number of observations underlying the cell-level aggregate, and the number of clusters from which these observations stem, vary across cells and indicators.

2.3 Control variables

We use three control variables. The first is population density (labelled population), which we compute based on the CIESIN Gridded Population of the World (GPWv4) dataset (Center for International Earth Science Information Network, 2016). We thereby use the values for 2010 for all years within our sample.[13]

The second control variable is the electrification rate (labelled electricity), which we derive from the DHS. Respondents are asked whether their household has electricity, without specifying the source. To calculate cell-level electrification rates, we use the same procedure applied to generate other cell-level development indicators.

The third control variable captures urbanization (labelled urban). DHS reports whether the clusters are urban or rural according to the classification used by the respective national statistical administration. For the small circular zones, the urban variable is simply an indicator that equals one if the cluster is classified as urban, and zero if it is classified as rural. For the PRIO-GRID cells, this variable is defined as the share of urban clusters among all DHS clusters whose center point falls within the cell.

Table 1 provides summary statistics for all our variables for both types of spatial units used in our analysis.

Table 1 around here

3 Empirical specification

Our main specification is a simple linear regression:

$$DHS_{ict} = \alpha_{ct} + \beta \ln(\text{light}_{ict} + 0.01) + \gamma \ln(\text{population}_{ict}) + \epsilon_{ict},$$

where DHS_{ict} represents one of our six development indicators in circular zone or PRIO-GRID cell i of country c in year t. We cluster standard errors ϵ_{ict} at the country level.

Three aspects of this specification are worth discussing. First, the country-year fixed effects α_{ct} imply that we look at the effect of nighttime lights to explain differences in our

[13] GPWv4 is mainly based on the 2010 round of censuses. It is not available for years prior to 2000. GPWv3, which is available for earlier years, has a much lower spatial resolution. Alternative population maps, like Landscan, which used nighttime lights as input in some years, are also not suitable for our purpose.

development indicators within single country-years or, equivalently, single surveys. Among others, we thereby account for differences across countries and waves in the survey items available to assess household wealth. In order to look at the association between nighttime lights and local human development across countries, we also run regressions in which we drop the country-year fixed effects. In these regressions, we add simple year fixed effects to absorb changes in the weather satellite technology and their sensor settings. We however exclude our indicators of household wealth in these regressions, as they are not comparable across countries.

Second, we control in all our regressions for population density. We could alternatively divide average nighttime lights by population to get a variable with the flavor of GDP per capita. However, controlling for population density is more flexible. In some regressions, we further control for electrification and urbanization. We do so because previous studies have shown that nighttime lights contain information about the size of urban clusters (Sutton et al., 2001; Sutton, 2003; Storeygard, 2016) as well as about electrification at the local level (Elvidge et al., 2011; Min et al., 2013). In this extended specification, in which we control for electrification and urbanization, the coefficient β measures the additional information contained in local nighttime lights that is not yet reflected in local population density, electrification and urbanization. This extended specification may however be less informative for researchers who are in need of a proxy for local human development and lack information on local electrification (and urbanization). Moreover, it is not suitable for comparing the explanatory power of nighttime lights, on the one hand, and electrification and urbanization, on the other hand. The reason is that our measures of electrification and urbanization are based on the same surveyed households as the development indicators, while the nighttime lights are the product of light emissions by all households in the respective spatial unit.

Third, we use the logarithm of nighttime lights and population density because the distributions of these variables are strongly right-skewed in both our samples. The log-transformed values give more weight to variation in nighttime lights and population density in the lower segments where observations are concentrated. Given the high share of spatial units with zero average nighttime lights in our samples, we follow Michalopoulos and Papaioannou (2013, 2014) and Hodler and Raschky (2014b,a) in adding a constant of 0.01 before taking the logarithm in order not to drop these observations from our two samples. This adjustment is in line with our assumption that zero values in the Stable Lights data do not reflect complete absence of light emissions. Normally, the coefficient of interest in a level-log regression can be interpreted as the increase in the dependent variable associated with a 100 percent increase in the log-transformed regressor. However, given that we add a small constant before taking logs, coefficient β only corresponds to an approximation to this effect. We however show that our results are robust when we refrain from adding a small constant before log-transforming average nighttime lights.

4 Empirical Findings

4.1 Main results

Table 2 shows the results from OLS regressions of our main specification. Panel A presents estimates based on circular zones around DHS clusters, and panel B estimates based on PRIO-GRID cells. Odd columns only control for population density, while even columns include electrification and urbanization as additional controls.

Table 2 around here

For all our development indicators and both types of spatial units, we find a positive and statistically highly significant association between nighttime lights and human development when controlling for population density only. The point estimates of the regressions based on circular zones imply that an increase in nighttime lights by 100 percent (which corresponds to a bit less than a within country-year standard deviation increase on the sample average) is associated with an increase in the mean household wealth value by 0.24, and by 0.19 for electricity-free wealth; an increase in primary school attendance by 2 in 100 children at primary school age; an increase in adults' years of education by 0.4 years; a drop in infant mortality by 1.2 deaths per 1000 live births; and an increase in the share of births assisted by professional health personnel by 4 in 100 births. The point estimates suggest stronger associations across PRIO-GRID cells.

When adding the local electrification and urbanization as controls, the magnitude of the coefficients on nighttime lights drops substantively for all our development indicators. This drop was to be expected, not least because these controls are based on the same surveyed households as the development indicators, while nighttime lights are not. Nevertheless, the coefficients on nighttime lights remain statistically significant for most indicators, except for years of schooling and infant mortality when using circular zones, and for infant mortality when using PRIO-GRID cells. Hence, while part of the variation in local nighttime lights are absorbed by variation in electrification and urbanization in the DHS clusters, nighttime lights still contain considerable information about local human development that goes beyond access to electricity and urbanization. Our results also confirm that higher rates of electrification and urbanization are positively associated with human development outcomes at the local level.

Overall, we conclude from Table 2 that variation in nighttime lights across small spatial units of different size within a country captures the relative human development level within that spatial unit. Nighttime lights that are relatively intense compared to the country average are indicative of a relatively wealthy, well-educated and healthy local population.

Next, we re-run the same regressions, but do not add a small constant before log-transforming the average nighttime lights. We thereby drop all spatial units where the

Stable Lights data report a DN of zero for all pixels. Results are reported in Table 3, which is organized like Table 2.

Table 3 around here

In spite of the substantially reduced sample, we still find statistically significant coefficients on nighttime light for all indicators when controlling for population density only. However, in the analysis based on circular zones around DHS clusters, the coefficients on nighttime light are no longer statistically significant when controlling for local electrification and urbanization. By contrast, when using the PRIO-GRID cells, coefficients on nighttime lights remain statistically significant for most development indicators even when adding these controls.

We further observe that the coefficients on nighttime lights tend to be larger for the small circular zones than for the PRIO-GRID cells when dropping spatial units where our light variable is zero. Assuming that nighttime lights and DHS-based development indicators are positively correlated at the very local level, we would indeed expect a stronger correlation between the two when restricting to nighttime lights data from the immediate surroundings of the location of the survey data collection. By contrast, in Table 2 where we included spatial units where our light variable is zero, we actually found that a higher spatial resolution returned weaker associations. A possible interpretation is that by including spatial units where our light variable is zero, we add additional noise to the relationship between nighttime lights and local development. For the circular zones, the share of spatial units without any nighttime lights is larger than for the PRIO-GRID cells; and so is the likelihood of not observing small variations in nighttime lights between circular zones, which may well have measurable differences in development. Hence, our results suggest that the most commonly used nighttime lights dataset, the Stable Lights series, may not appropriately capture differences in local development in places with very low nighttime light emissions.[14]

We next examine whether nighttime lights are also positively associated with local human development when we compare our subnational spatial units not only within countries, but across the African continent. We therefore estimate our main specification after replacing the country-year fixed effects by year fixed effects. Given that the methodology used to derive our wealth indicators is specific to country and survey waves, we exclude them from this part of the analysis. Results are presented in Table 4, which is organized as previous tables.

Table 4 around here

[14]Table A.2 in Appendix A takes a closer look at the characteristics of spatial units in our sample which contain only Stable Lights pixels with DN zero. It documents that these spatial units tend to have limited access to electricity and to be rural and sparsely populated. Also, they are more common in poorer countries.

We find that local nighttime lights are again positively correlated with education and health at the local level. Effect sizes are very similar as in Table 2, except for infant mortality, where the analysis across countries yields a substantially stronger association. Moreover, the cross-country association between nighttime lights and local human development is even more robust to controlling for electrification and urbanization than the corresponding within-country association. We conclude that nighttime lights are also a good proxy for local human development when using a research design that exploits cross-country variation between small spatial units.[15]

4.2 Robustness

We run two robustness tests for our main results. First, we check whether the relationship between nighttime lights and local human development holds also when we account for the substantial variation in the numbers of observations that underlie our local development indicators. For the circular zones around DHS clusters, this variation is driven mainly by differences in the number of households sampled per cluster. For the PRIO-GRID cells, the number of observations which feed into cell-level aggregates is also driven by the number of clusters that fall within the cell. We tend to have more clusters per cell in densely populated areas, and few or only a single cluster per cell in remote rural areas. We therefore replicate our main estimates as weighted regressions, where we apply as weights the number of observations used to compute each development indicator value for a given zone. We thereby give more weight to spatial zones for which we can be more confident that our development indicators capture the actual level of development on the ground. Results are presented in Table A.3 in Appendix A. When using the circular zones around DHS clusters, the weighted regressions yield very similar results as the main regressions presented in Table 2. When using the PRIO-GRID cells, the weighted regressions return the same overall pattern of results as the unweighted regressions, except that the coefficients on nighttime lights become statistically insignificant in some regressions in which we control for local electrification and urbanization.

Second, we have so far ignored spatial overlaps of circular zones around DHS clusters. In our main regressions, we treat each zone as an independent observation, no matter

[15]We have also examined the association between changes in nighttime lights and changes in human development within grid cells over time. We have done so for PRIO-GRID cells as well as for considerably smaller cells. We discuss our methodology and our results in Appendix B. We do not find any statistically significant associations between changes in nighttime lights and changes in DHS-based human development indicators. It is hard to interpret this absence of a measurable association. One interpretation could be that changes in nighttime lights over time are truly a bad proxy for changes in human development at the subnational level. A second interpretation could be that DHS are unsuitable to capture changes in local development over time as they are not designed as a panel, so that changes in our development indicators are based on comparisons of different households from different villages or urban blocks. Given that Hodler and Raschky (2014b) document a relation between changes in nighttime lights and province-level GDP per capita over time, while we find not even a positive and statistically significant association between changes in nighttime lights and our DHS-based wealth measures, we tend to favor the second interpretation.

whether it is located in isolation or whether it is part of an agglomeration of settlements where several clusters were sampled. We now take into account whether a circular zone overlaps with other circular zones of the same survey wave. We do this by again applying weights to our main regressions based on circular zones around DHS clusters.[16] This time we use as weights $w_i = \frac{1}{1+Z_i}$, where Z_i is the number of zones from the same survey wave that overlap with a given zone i. We now give more weight to spatial zones located in isolation, and less weight to zones located in an area where we have information from several nearby DHS clusters. As shown in Table A.4 in Appendix A, results are very similar as our main results presented in Table 2.

We conclude from these robustness checks that our results are by and large robust to variations in the number of observations available to calculate local development indicators, and to variation in the number of survey clusters sampled in a particular area.

4.3 Heterogeneity

So far we have shown average correlations between nighttime lights and different development indicators across a wide range of economic, political and geographic contexts. We now check whether the association between nighttime lights and local development differs depending on the country's economic and political context as well as local settlement characteristics. We thereby focus on the association between nighttime lights and household wealth, as most researchers may still be primarily interested in the usefulness of nighttime lights as a proxy for economic development.

First, we use three measures that capture the economic and political country context, namely GDP per capita, the share of agricultural value added in GDP, and the level of democracy measured by the Polity2 score.[17] For each of these country variables, we obtain the respective values for each country for each survey year. We derive quartiles in the distribution of these country variables in a sample of country-survey years. For each country variable, we then use these quartiles to split our two samples of spatial units into four sub-samples each. We re-run our main specification of household wealth on nighttime lights and population density within countries, i.e., the same specification as in Table 2, column (1), on these sub-samples.

Results are plotted in Figure 2, with dots indicating point estimates and bars indicating 95% confidence intervals from the regressions on the respective sub-samples. In each plot, the first estimate is from the sub-sample below the first quartile; the second dot from the sub-sample between the first and the second quartile; and so forth. Column (a) shows results for our circular zones around DHS clusters, and column (b) for the PRIO-GRID

[16]Spatial overlaps are not an issue for our analysis based on grid cells, where the procedure applied to derive cell-level aggregate values already accounts for overlaps between buffer zones around DHS clusters.

[17]More specifically, we use gross domestic product (GDP) per capita in constant 2005 dollars (World Bank, 2016), value added in agriculture (including forestry, hunting, fishing, cultivation of crops, and livestock production) as a share of GDP (World Bank, 2016), and the Polity2 score (Marshall et al., 2015).

cells.

Figure 2 around here

The association between nighttime lights and household wealth tends to become stronger as the country's level of GDP per capita rises, implying that nighttime lights might be a better proxy for local economic development in richer African countries. The effects of the agriculture share and the level of democracy on the association between nighttime lights and household wealth seem to be U-shaped.

We now check whether the association between nighttime lights and local human development changes with the local settlement patterns, and with local electrification rates. Specifically, we disaggregate our samples by the local population density in 2010; by the local share of households with electricity access as measured by the DHS; and by urbanization (which is measured by the urban indicator for our circular zones and as the share of clusters classified as urban for the PRIO-GRID cells). We take the distribution of each of these indicators in our samples of circular zones and PRIO-GRID cells, respectively, and split the samples by quartiles. Again, we re-run our main regression on these sub-samples.

Results are shown in Figure 3, which is organized analogous to Figure 2.

Figure 3 around here

The association between nighttime lights and household wealth tends to be stronger in areas that are less densely populated, but that have higher rates of electrification. It is however similar in urban and rural contexts.

5 Conclusions

Nighttime lights, calculated from weather satellite recordings, are increasingly used by social scientists as a proxy for economic activity or development in regions for which disaggregated data from statistical offices are not available. Previous literature has shown that nighttime light intensity correlates with GDP at the level of countries, provinces, and grid cells. This paper complements the evidence underpinning the use of nighttime light in the social sciences by two important aspects: First, we have studied whether nighttime lights are an accurate proxy for development at a higher spatial resolution. We confirm that nighttime lights can serve as a proxy for development at the very local level, such as municipalities, as well as across PRIO-GRID cells. This finding may encourage social scientists to consider nighttime lights as a proxy for development in studies requiring very high spatial resolution or relying on PRIO-GRID data. They should however carefully consider whether to keep or to drop locations which contain only Stable Lights pixels with a value of zero. Dropping these apparently dark locations has the disadvantage that the sample size may decrease substantively (e.g., by 51 and 42 percent in our samples of small

circular zones and PRIO-GRID cells, respectively), but the advantage that nighttime lights become an ever better proxy for development for those locations that remain in the sample.

Second, we have examined whether nighttime lights can capture not only economic development, but also other dimensions of human development at the local level. We can indeed confirm that nighttime lights capture human development as measured in household wealth, education and health. This implies that future studies concerned with subnational differences in human development may well use nighttime lights in absence of other reliable disaggregated data. It also implies that some of the existing studies that used nighttime lights to study, e.g., comparative development merit broader interpretations, considering that differences in nighttime lights across subnational regions reveal not only differences in economic activity, but also differences in the level of human development of the local population.

References

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016): "Ethnic inequality," *Journal of Political Economy*, 124, 428–488.

Baugh, K., C. Elvidge, T. Ghosh, and D. Ziskin (2013): "Development of a 2009 Stable Lights product using DMSP-OLS data," *Proceedings of the Asia-Pacific Advanced Network*, 30, 114–130.

Besley, T. and M. Reynal-Querol (2014): "The legacy of historical conflict: Evidence from Africa," *American Political Science Review*, 108, 319–336.

Brown, M., B. Guin, and K. Kirschenmann (2016): "Microfinance banks and financial inclusion," *Review of Finance*, 20, 907–946.

Center for International Earth Science Information Network (2016): *Gridded population of the world, Version 4 (GPWv4)*, Palisades, NY, USA: Center for International Earth Science Information Network (CIESIN), Columbia University.

Chen, X. (2015): "Explaining subnational infant mortality and poverty rates: What can we learn from nightime lights?" *Spatial Demography*, 3, 27–53.

Chen, X. and W. Nordhaus (2011): "Using luminosity data as a proxy for economic statistics," *Proceedings of the National Academy of Sciences of the United States of America*, 108, 8589–8594.

De Luca, G., R. Hodler, P. Raschky, and M. Valsecchi (2015): *Ethnic favoritism: An axiom of politics?*, CESifo Working Paper 5209.

Doll, C., J.-P. Muller, and J. Morley (2006): "Mapping regional economic activity from nighttime light satellite imagery," *Ecological Economics*, 57, 75–92.

Donaldson, D. and A. Storeygard (2016): "The view from above: Applications of satellite data in economics," *Journal of Economic Perspectives*, 30, 171–198.

Elvidge, C., K. Baugh, P. Sutton, B. Bhaduri, B. Tuttle, T. Ghosh, D. Ziskin, and E. Erwin (2011): "Who's in the dark: Satellite based estimates of electrification rates," in *Urban Remote Sensing*, ed. by X. Yang, Hoboken, NJ, USA: Wiley, 211–224.

Elvidge, C., P. Sutton, T. Ghosh, B. Tuttle, K. Baugh, B. Bhaduri, and E. Bright (2009): "A global poverty map derived from satellite data," *Computers & Geosciences*, 35, 1652–1660.

Filmer, D. and K. Scott (2012): "Assessing asset indices," *Demography*, 49, 359–392.

Gennaioli, N., R. La Porta, F. Lopez De Silanes, and A. Shleifer (2014): "Growth in regions," *Journal of Economic Growth*, 19, 259–309.

Ghosh, T., S. Anderson, R. Powell, L. Powell, P. Sutton, and C. Elvidge (2009): "Estimation of Mexico's informal economy and remittances using nighttime imagery," *Remote Sensing*, 1, 418–444.

Henderson, J. V., T. Squires, A. Storeygard, and D. Weil (2017): "The global distribution of economic activity: Nature, history, and the role of trade," *Quarterly Journal of Economics*, forthcoming.

Henderson, J. V., A. Storeygard, and D. Weil (2012): "Measuring economic growth from outer space," *American Economic Review*, 102, 994–1028.

Hodler, R. and P. Raschky (2014a): "Economic shocks and civil conflict at the regional level," *Economics Letters*, 124, 530–533.

——— (2014b): "Regional favoritism," *Quarterly Journal of Economics*, 129, 995–1033.

ICF International, ed. (1992-2013): *Demographic and Health Surveys (various) [Datasets]*, Calverton, MD, USA: ICF International.

Jean, N., M. Burke, M. Xie, W. Davis, D. Lobell, and S. Ermon (2016): "Combining satellite imagery and machine learning to predict poverty," *Science*, 353, 790–794.

Letu, H., M. Hara, G. Tana, and F. Nishio (2012): "A saturated light correction method for DMSP/OLS nighttime satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, 50, 389–396.

Marshall, M., K. Jaggers, and T. Gurr (2015): *Polity IV project [Datasets]*.

Mellander, C., J. Lobo, K. Stolarick, and Z. Matheson (2015): "Nighttime light data: A good proxy measure for economic activity?" *PloS one*, 10, 1–18.

Michalopoulos, S. and E. Papaioannou (2013): "Pre-colonial ethnic institutions and contemporary African development," *Econometrica*, 81, 113–152.

——— (2014): "National institutions and subnational development in Africa," *Quarterly Journal of Economics*, 129, 151–213.

Min, B., K. M. Gaba, O. F. Sarr, and A. Agalassou (2013): "Detection of rural electrification in Africa using DMSP-OLS night lights imagery," *International Journal of Remote Sensing*, 34, 8118–8141.

Montalvo, J. and M. Reynal-Querol (2016): *Ethnic diversity and growth: Revisiting the evidence*, Mimeo.

Noor, A., V. Alegana, P. Gething, A. Tatem, and R. Snow (2008): "Using remotely sensed nighttime light as a proxy for poverty in Africa," *Population Health Metrics*, 6, 5.

Rutstein, S. and K. Johnson (2004): *The DHS wealth index: DHS Comparative Reports No. 6*, Calverton, MD, USA: ORC Macro.

Sen, A., J. Stiglitz, and J. P. Fitoussi (2010): *Mismeasuring our lives: Why GDP doesn't add up*, New York, NY, USA: New Press.

Storeygard, A. (2016): "Farther on down the road: Transport costs, trade and urban growth in Sub-Saharan Africa," *Review of Economic Studies*, 83, 1263–1295.

Sutton, P. (2003): "A scale-adjusted measure of urban sprawl using nighttime satellite imagery," *Remote Sensing of Environment*, 86, 353–369.

Sutton, P. and R. Costanza (2002): "Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation," *Ecological Economics*, 41, 509–527.

Sutton, P., C. Elvidge, and T. Ghosh (2007): "Estimation of gross domestic product at sub-national scales using nighttime satellite imagery," *International Journal of Ecological Economics & Statistics*, 8, 5–21.

Sutton, P., D. Roberts, C. Elvidge, and K. Baugh (2001): "Census from heaven: An estimate of the global human population using nighttime satellite imagery," *International Journal of Remote Sensing*, 22, 3061–3076.

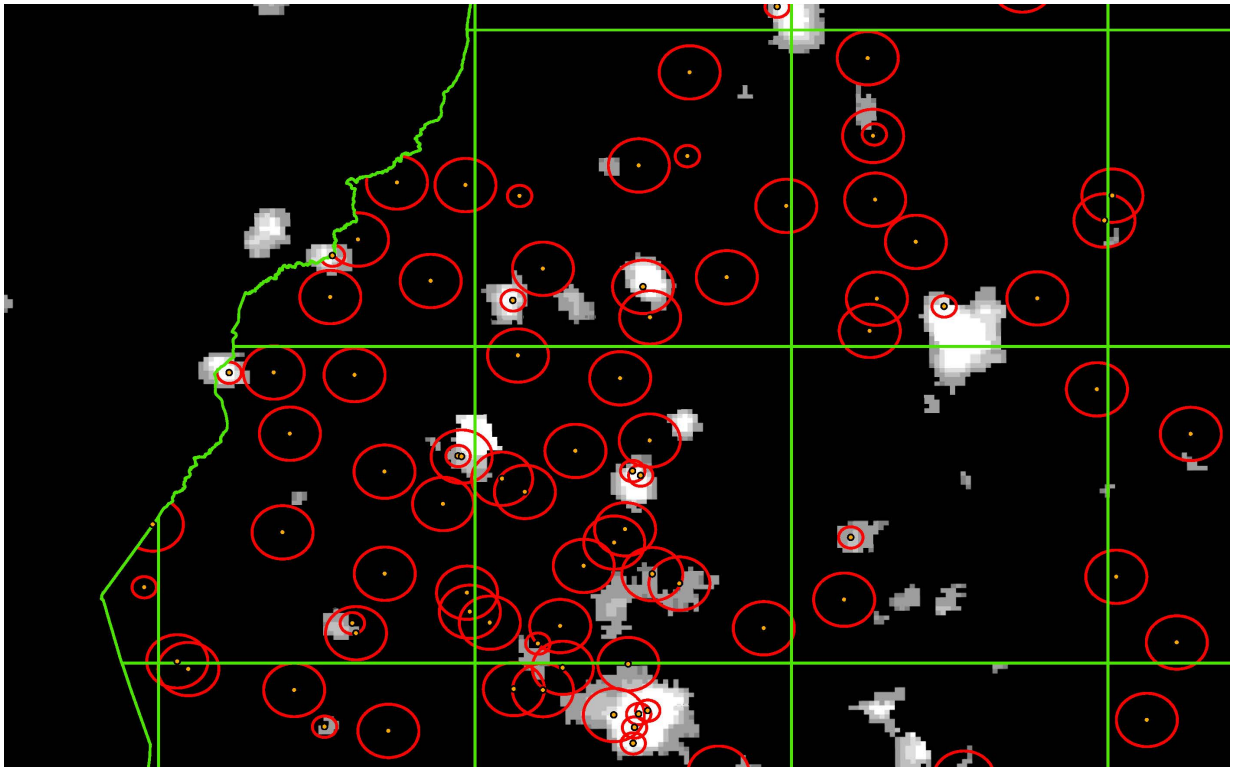
Tollefsen, A. F., H. Strand, and H. Buhaug (2012): "PRIO-GRID: A unified spatial data structure," *Journal of Peace Research*, 49, 363–374.

Weidmann, N. and S. Schutte (2017): "Using night light emissions for the prediction of local wealth," *Journal of Peace Research*, 54, 125–140.

World Bank (2016): *World Development Indicators 2016*, Washington, DC, USA: World Bank.

Figures and Tables

Figure 1: Illustration of DHS clusters, nighttime lights and our spatial units for Western Kenya



Notes: Dots represent reported DHS cluster center points from Kenya DHS 2008. Circular zones around these center points (in red) have a 2 km radius for urban clusters and a 5 km radius for rural clusters. The grid cells of 0.5×0.5 decimal degrees (in green) are aligned to the PRIO-GRID. Nighttime lights underlay the map, with brighter tones of grey implying more intense nighttime lights.

Table 1: Summary statistics

	Obs.	Mean	Std. Dev. overall/within	Min.	Max.
Panel A: Small circular zones					
light	29,169	11.75	19.56/14.78	0	63
household wealth	26,759	2.953	1.147/1.068	1	5
e-free wealth	26,702	2.869	1.161/1.076	1	5
school attendance	28,788	0.722	0.279/0.255	0	1
years of schooling	28,860	4.751	3.159/2.797	0	19
infant mortality	28,818	58.91	83.74/75.37	0	1,000
birth assistance	27,901	0.565	0.349/0.315	0	1
population	28,920	1,491	4,871/2,984	0	80,956
electricity	29,051	0.356	0.424/0.343	0	1
urban	29,169	0.322	0.467/0.417	0	1
Panel B: PRIO-GRID cells					
light	7,515	0.775	3.377/3.132	0	57.69
household wealth	7,144	2.585	0.907/0.825	1	5
e-free wealth	7,275	2.717	0.950/0.882	1	5
school attendance	7,436	0.613	0.287/0.263	0	1
years of schooling	7,442	3.796	2.657/2.363	0	12.73
infant mortality	7,498	65.63	60.81/57.47	0	571.4
birth assistance	7,123	0.428	0.304/0.279	0	1
population	7,515	115.2	342.8/307.9	0	7,153
electricity	7,500	0.199	0.305/0.278	0	1
urban	7,515	0.191	0.298/0.273	0	1

Notes: Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in panel A, and PRIO-GRID cells in panel B. All variables are described in the main text. Reported standard deviations are the overall standard deviation and the standard deviation within country-years.

Table 2: Nighttime lights and development within country-years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep. var.:	household wealth		e-free wealth		school attendance		years of schooling		infant mortality		birth assistance	
Panel A: Small circular zones												
ln(light+0.01)	0.238***	0.056***	0.187***	0.111***	0.023***	0.005**	0.416***	0.052	-1.213***	0.237	0.044***	0.011***
	(0.009)	(0.014)	(0.028)	(0.038)	(0.004)	(0.002)	(0.023)	(0.039)	(0.342)	(0.382)	(0.003)	(0.003)
ln(population)	0.094***	0.067***	-0.005	-0.012	0.017***	0.015**	0.242***	0.183***	-0.536	-0.342	0.022***	0.017***
	(0.033)	(0.022)	(0.047)	(0.037)	(0.005)	(0.005)	(0.041)	(0.030)	(0.495)	(0.522)	(0.006)	(0.005)
electricity		1.547***		1.123***		0.177***		3.101***		-15.897***		0.259***
		(0.117)		(0.328)		(0.038)		(0.270)		(3.420)		(0.031)
urban		0.761***		-0.037		0.054***		1.556***		-3.579		0.157***
		(0.181)		(0.384)		(0.014)		(0.510)		(2.126)		(0.031)
Observations	26,492	26,492	26,435	26,435	28,485	28,374	28,554	28,438	28,516	28,516	27,605	27,605
R^2	0.537	0.708	0.246	0.290	0.500	0.526	0.615	0.705	0.050	0.052	0.431	0.492
Panel B: PRIO-GRID cells												
ln(light+0.01)	0.346***	0.085***	0.268***	0.117***	0.039***	0.008**	0.612***	0.166***	-2.479***	-0.728	0.064***	0.014**
	(0.026)	(0.014)	(0.042)	(0.023)	(0.005)	(0.003)	(0.053)	(0.039)	(0.563)	(0.760)	(0.006)	(0.005)
ln(population)	-0.041	0.031	-0.051	-0.009	0.013	0.022*	-0.039	0.086	2.163*	1.669	-0.006	0.008
	(0.076)	(0.046)	(0.067)	(0.046)	(0.011)	(0.011)	(0.101)	(0.070)	(1.116)	(1.152)	(0.011)	(0.007)
electricity		2.039***		1.213***		0.250***		3.558***		-14.047**		0.377***
		(0.110)		(0.394)		(0.060)		(0.279)		(5.593)		(0.036)
urban		0.726***		0.390*		0.087***		1.228***		-4.762		0.153***
		(0.108)		(0.193)		(0.019)		(0.284)		(3.373)		(0.025)
Observations	7,133	7,133	7,264	7,264	7,422	7,410	7,428	7,415	7,484	7,484	7,109	7,109
R^2	0.360	0.644	0.265	0.352	0.518	0.559	0.628	0.726	0.104	0.107	0.459	0.552

Notes: Linear regressions with country-year fixed effects on a sample including all geo-coded DHS in African countries from 1992-2013. Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in panel A, and PRIO-GRID cells in panel B. All variables are described in the main text. Standard errors are clustered at the country level. ***, **, * indicate significance at the 1, 5 and 10%-level, respectively.

Table 3: Nighttime lights and development within country-years, excluding dark spatial units

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep. var.:	household wealth		e-free wealth		school attendance		years of schooling		infant mortality		birth assistance	
Panel A: Small circular zones												
ln(light)	0.389***	0.055	0.268***	0.214***	0.033***	0.005	0.787***	0.054	-2.598***	0.212	0.068***	0.007
	(0.010)	(0.043)	(0.048)	(0.076)	(0.005)	(0.003)	(0.059)	(0.098)	(0.837)	(1.134)	(0.006)	(0.007)
ln(population)	0.038	0.062***	-0.070	-0.072	0.004*	0.006**	0.145***	0.189***	-0.341	-0.506	0.009***	0.013***
	(0.022)	(0.019)	(0.046)	(0.044)	(0.002)	(0.002)	(0.024)	(0.021)	(0.682)	(0.719)	(0.003)	(0.003)
electricity		1.385***		1.277***		0.172***		3.051***		-17.597***		0.230***
		(0.131)		(0.272)		(0.025)		(0.254)		(3.523)		(0.030)
urban		0.805***		-0.518		0.033**		1.754**		-3.155		0.158***
		(0.277)		(0.448)		(0.016)		(0.685)		(3.336)		(0.045)
Observations	12,498	12,498	12,445	12,445	13,926	13,867	13,982	13,920	13,856	13,856	13,660	13,660
R^2	0.550	0.731	0.370	0.443	0.339	0.379	0.475	0.585	0.039	0.041	0.299	0.375
Panel B: PRIO-GRID cells												
ln(light)	0.310***	0.070***	0.228***	0.095***	0.037***	0.008**	0.559***	0.130**	-1.591**	-0.123	0.058***	0.012**
	(0.021)	(0.011)	(0.038)	(0.022)	(0.006)	(0.003)	(0.060)	(0.050)	(0.602)	(1.079)	(0.006)	(0.005)
ln(population)	-0.071	0.012	-0.090	-0.043	0.002	0.012	-0.039	0.113	1.278	0.760	-0.008	0.008
	(0.074)	(0.034)	(0.059)	(0.035)	(0.009)	(0.008)	(0.119)	(0.071)	(1.280)	(1.180)	(0.011)	(0.006)
electricity		2.006***		1.250***		0.282***		3.697***		-12.644**		0.353***
		(0.120)		(0.360)		(0.053)		(0.272)		(5.587)		(0.039)
urban		0.698***		0.255		0.047**		1.164***		-3.964		0.160***
		(0.131)		(0.262)		(0.017)		(0.334)		(4.407)		(0.027)
Observations	4,100	4,100	4,154	4,154	4,308	4,304	4,309	4,305	4,340	4,340	4,139	4,139
R^2	0.338	0.716	0.203	0.324	0.496	0.563	0.581	0.717	0.102	0.106	0.463	0.587

Notes: Linear regressions with country-year fixed effects on a sample including all geo-coded DHS in African countries from 1992-2013. Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in panel A, and PRIO-GRID cells in panel B. All variables are described in the main text. Standard errors are clustered at the country level. ***, **, * indicate significance at the 1, 5 and 10%-level, respectively.

Table 4: Nighttime lights and development across Africa

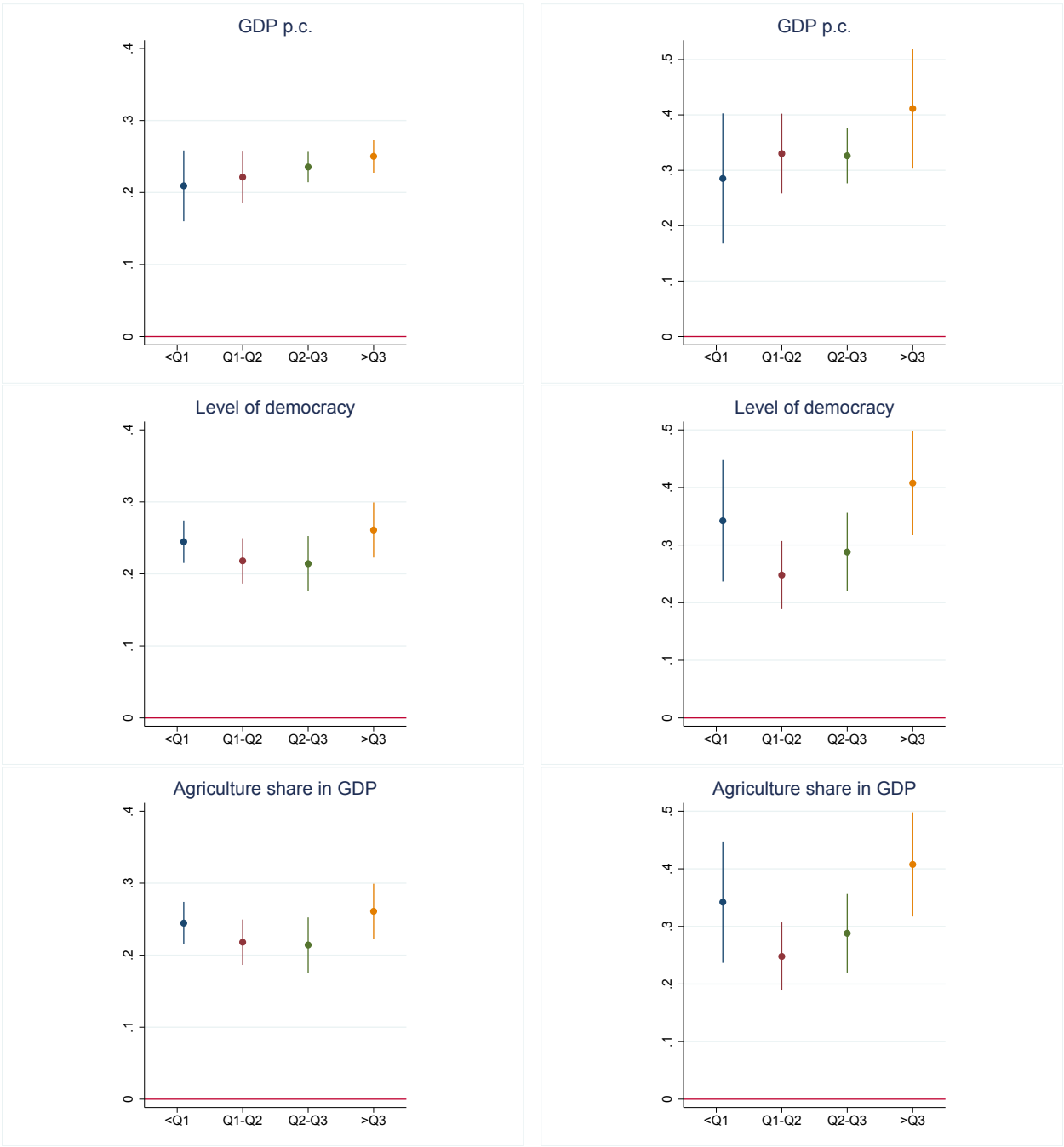
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.:	school attendance		years of schooling		infant mortality		birth assistance	
Panel A: Small circular zones								
ln(light+0.01)	0.025***	0.007*	0.467***	0.095*	-2.827***	-0.962*	0.047***	0.016***
	(0.006)	(0.004)	(0.049)	(0.051)	(0.538)	(0.480)	(0.004)	(0.005)
ln(population)	0.020**	0.020**	0.170	0.150	0.517	0.445	0.011	0.009
	(0.009)	(0.009)	(0.112)	(0.101)	(0.923)	(0.920)	(0.012)	(0.011)
electricity		0.180***		2.937***		-24.057***		0.212***
		(0.031)		(0.309)		(4.565)		(0.034)
urban		0.028**		1.455***		3.473*		0.160***
		(0.013)		(0.524)		(1.920)		(0.029)
Observations	28,485	28,374	28,554	28,438	28,516	28,516	27,605	27,605
R^2	0.286	0.310	0.432	0.519	0.027	0.032	0.275	0.331
Panel B: PRIO-GRID cells								
ln(light+0.01)	0.056***	0.023**	0.746***	0.286***	-4.910***	-2.920***	0.079***	0.026***
	(0.009)	(0.009)	(0.076)	(0.080)	(0.886)	(0.860)	(0.008)	(0.009)
ln(population)	-0.014	-0.001	-0.301**	-0.140	5.719***	4.860***	-0.047**	-0.029
	(0.014)	(0.014)	(0.143)	(0.127)	(0.934)	(0.993)	(0.020)	(0.018)
electricity		0.233***		2.940***		-17.437***		0.323***
		(0.054)		(0.394)		(4.949)		(0.044)
urban		0.064**		1.218***		1.733		0.165***
		(0.027)		(0.347)		(4.909)		(0.029)
Observations	7,422	7,410	7,428	7,415	7,484	7,484	7,109	7,109
R^2	0.275	0.311	0.397	0.480	0.069	0.072	0.242	0.331

Notes: Linear regressions with year fixed effects on a sample including all geo-coded DHS in African countries from 1992-2013. Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in panel A, and PRIO-GRID cells in panel B. All variables are described in the main text. Standard errors are clustered at the country level. ***, **, * indicate significance at the 1, 5 and 10%-level, respectively.

Figure 2: Nighttime lights and household wealth for different country characteristics

(a) Small circular zones

(b) PRIO-GRID cells

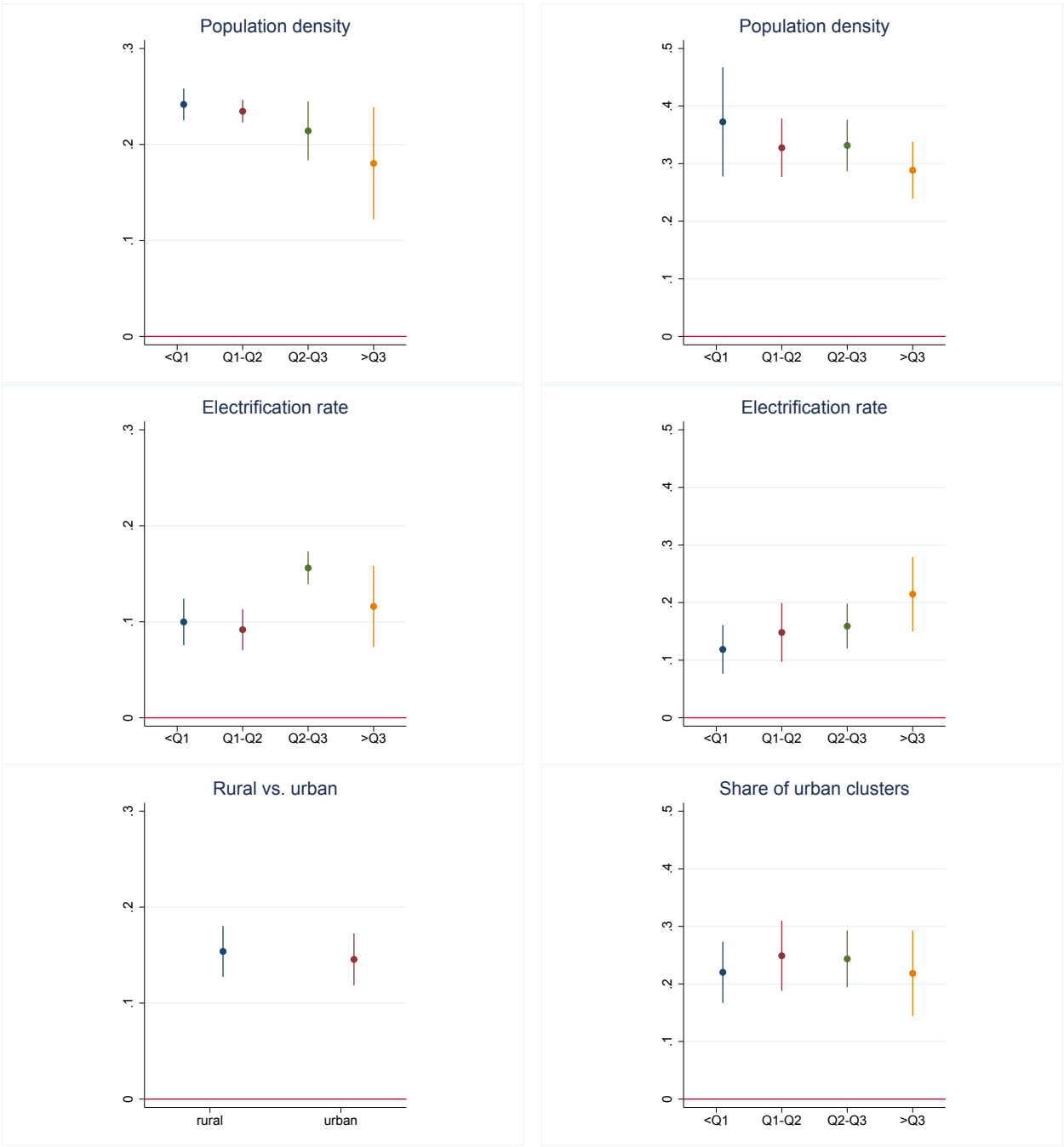


Notes: Figures show the relation between nighttime lights and household wealth in different sub-samples. Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in column (a), and PRIO-GRID cells in column (b). Each dot shows the point estimate, and each bar the 95% confidence interval for the coefficient on ln(light+0.01) in a regression of wealth on ln(light+0.01), ln(population) and country-year fixed effects (as in Table 2, column (1)). Regressions are run on sub-samples drawn by quartiles in the distribution of country's GDP per capita, level of democracy (Polity2 score), or share of agricultural value added in GDP, respectively. Q1 indicates the sub-sample below the first quartile, Q2 the sub-sample between the 1st and the 2nd quartile, etc.

Figure 3: Nighttime lights and household wealth for different local characteristics

(a) Small circular zones

(b) PRIO-GRID cells



Notes: Figures show the relation between nighttime lights and household wealth in different sub-samples. Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in column (a) and PRIO-GRID cells in column (b). Each dot shows the point estimate, and each bar the 95% confidence interval for the coefficient on ln(light+0.01) in a regression of wealth on ln(light+0.01), ln(population) and country-year fixed effects (as in Table 2, column (1)). Regressions are run on sub-samples drawn by quartiles in the distribution of the local of population density, the local electrification rate, and the local urbanization (replace by an indicator in column (a)), respectively. Q1 indicates the sub-sample below the first quartile, Q2 the sub-sample between the 1st and the 2nd quartile, etc.

Appendix A: Additional tables

Table A.1: List of countries and DHS waves in our sample

County	DHS waves
Benin	1996, 2001, 2011
Burkina Faso	1998, 2003, 2010
Burundi	2010
Cameroon	2004, 2011
Cote d'Ivoire	1994, 1998, 2011
Dem. Rep. Congo	2007, 2013
Egypt	1992, 1995, 2000, 2005, 2008
Ethiopia	2000, 2005, 2010
Ghana	1998, 2003, 2008
Guinea	1999, 2005, 2012
Kenya	2003, 2008
Lesotho	2004, 2009
Liberia	2006, 2013
Madagascar	1997, 2008
Malawi	2000, 2004, 2010
Mali	1995, 2001, 2006, 2012
Mozambique	2011
Niger	1998
Nigeria	2003, 2008, 2011
Namibia	2000, 2006, 2013
Rwanda	2005, 2010
Sierra Leone	2008, 2013
Senegal	1997, 2005, 2010
Swaziland	2006
Togo	1998, 2013
Tanzania	1999, 2009
Uganda	2000, 2006, 2011
Zambia	2007, 2013
Zimbabwe	1999, 2005, 2010

Table A.2: Determinants of non-zero nighttime lights

	(1)	(2)	(3)
Dep. var.: Indicator for non-zero nighttime lights			
Panel A: Small circular zones			
ln(population)	0.057***	0.064***	0.068***
	(0.010)	(0.012)	(0.013)
electricity	0.576***	0.471***	0.464***
	(0.069)	(0.073)	(0.084)
urban	0.184***	0.205***	0.202***
	(0.061)	(0.061)	(0.064)
ln(GDP pc)		0.097***	
		(0.018)	
Observations	28,746	28,746	28,746
R^2	0.559	0.571	0.590
Panel B: PRIO-GRID cells			
ln(population)	0.082***	0.112***	0.111***
	(0.018)	(0.017)	(0.020)
electricity	0.501***	0.274***	0.266***
	(0.064)	(0.058)	(0.081)
urban	0.227***	0.274***	0.297***
	(0.049)	(0.050)	(0.047)
ln(GDP pc)		0.171***	
		(0.031)	
Observations	7,486	7,486	7,486
R^2	0.289	0.326	0.363
Year fixed effects	YES	YES	NO
Country-year fixed effects	NO	NO	YES

Notes: Linear fixed effects regressions on a sample including all geo-coded DHS in African countries from 1992–2013. The dependent variable is an indicator variable that equals zero (one) if the mean light value in the zone is zero (above zero). Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in panel A, and PRIO-GRID cells in panel B. Standard errors are clustered at the country level. ***, **, * indicate significance at the 1, 5 and 10%-level, respectively.

Table A.3: Nighttime lights and development within country-years, weighted by number of observations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep. var.:	household wealth		e-free wealth		school attendance		years of schooling		infant mortality		birth assistance	
Panel A: Small circular zones												
ln(light+0.01)	0.235***	0.056***	0.182***	0.083***	0.023***	0.003*	0.422***	0.065*	-1.305***	0.179	0.044***	0.013***
	(0.009)	(0.012)	(0.020)	(0.024)	(0.003)	(0.002)	(0.025)	(0.035)	(0.299)	(0.239)	(0.004)	(0.003)
ln(population)	0.085***	0.060***	0.039	0.026	0.021***	0.017***	0.245***	0.177***	-0.522	-0.251	0.027***	0.021***
	(0.029)	(0.017)	(0.033)	(0.027)	(0.005)	(0.005)	(0.039)	(0.029)	(0.402)	(0.433)	(0.006)	(0.005)
electricity		1.562***		1.111***		0.209***		3.089***		-14.909***		0.266***
		(0.102)		(0.198)		(0.038)		(0.259)		(2.960)		(0.028)
urban		0.745***		0.222		0.064***		1.490***		-5.738***		0.158***
		(0.146)		(0.236)		(0.017)		(0.444)		(1.829)		(0.025)
Observations	26,492	26,492	26,435	26,435	28,485	28,374	28,554	28,438	28,505	28,505	27,605	27,605
R^2	0.555	0.728	0.330	0.398	0.502	0.535	0.630	0.716	0.064	0.066	0.454	0.514
Panel B: PRIO-GRID cells												
ln(light+0.01)	0.422***	0.068*	0.394***	0.221***	0.046***	0.009	0.797***	0.068	-2.178***	0.274	0.076***	0.011
	(0.025)	(0.037)	(0.058)	(0.075)	(0.007)	(0.005)	(0.058)	(0.073)	(0.651)	(0.878)	(0.008)	(0.009)
ln(population)	-0.063	0.024	-0.196*	-0.141*	-0.002	0.005	-0.024	0.139***	-0.279	-0.623	-0.000	0.008
	(0.039)	(0.028)	(0.098)	(0.075)	(0.009)	(0.006)	(0.057)	(0.035)	(0.860)	(0.985)	(0.011)	(0.005)
electricity		1.323***		1.775***		0.178***		2.804***		-6.687		0.248***
		(0.247)		(0.476)		(0.048)		(0.464)		(6.646)		(0.063)
urban		1.366***		-0.471		0.128***		2.858***		-13.418***		0.302***
		(0.388)		(0.721)		(0.042)		(0.679)		(3.451)		(0.067)
Observations	7,133	7,133	7,264	7,264	7,422	7,410	7,428	7,415	7,484	7,484	7,109	7,109
R^2	0.676	0.859	0.535	0.601	0.637	0.682	0.743	0.863	0.267	0.277	0.626	0.726

Notes: Weighted linear regressions with country-year fixed effects on a sample including all geo-coded DHS in African countries from 1992-2013. Weights are number of observations from the DHS data from which local development indicators are computed for each spatial unit. Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters in panel A, and PRIO-GRID cells in panel B. All variables are described in the main text. Standard errors are clustered at the country level. ***, **, * indicate significance at the 1, 5 and 10%-level, respectively.

Table A.4: Nighttime lights and development within country-years, weighted by inverse number of overlapping circular zones

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep. var.:	household wealth		e-free wealth		school attendance		years of schooling		infant mortality		birth assistance	
	Small circular zones											
ln(light+0.01)	0.242***	0.062***	0.164***	0.072***	0.023***	0.004**	0.418***	0.080***	-1.213***	-0.050	0.046***	0.011***
	(0.010)	(0.008)	(0.031)	(0.019)	(0.004)	(0.002)	(0.028)	(0.024)	(0.408)	(0.360)	(0.003)	(0.003)
ln(population)	0.077*	0.066**	0.029	0.025	0.023***	0.022***	0.202***	0.182***	-0.027	0.016	0.023***	0.020***
	(0.043)	(0.028)	(0.039)	(0.031)	(0.007)	(0.008)	(0.057)	(0.047)	(0.576)	(0.628)	(0.007)	(0.006)
electricity		1.687***		0.996***		0.190***		3.208***		-14.311***		0.292***
		(0.092)		(0.313)		(0.043)		(0.238)		(4.340)		(0.029)
urban		0.641***		0.219		0.059***		1.191***		-1.712		0.146***
		(0.107)		(0.206)		(0.016)		(0.325)		(2.295)		(0.024)
Observations	26,492	26,492	26,435	26,435	28,485	28,374	28,554	28,438	28,516	28,516	27,605	27,605
R^2	0.474	0.668	0.217	0.269	0.499	0.526	0.624	0.710	0.049	0.051	0.410	0.476

Notes: Weighted linear regressions with country-year fixed effects on a sample including all geo-coded DHS in African countries from 1992-2013. Units of observation are circular zones of 2 km (5 km) radius around urban (rural) DHS clusters, and weights are the inverse number of overlaps with other circular zones. All variables are described in the main text. Standard errors are clustered at the country level. ***, **, * indicate significance at the 1, 5 and 10%-level, respectively.

Appendix B: Analysis over time

To extend upon our main analysis, we examine the association between changes in nighttime lights and changes in human development within spatial units over time. In addition to the 0.5×0.5 decimal degree PRIO-GRID cells, we now also use a higher resolution grid of 0.1×0.1 decimal degree cells. With a size of approximately 10 km $\times$ 10 km at the equator, they correspond more closely to the smaller spatial units in our cross-sectional analysis. For both grid sizes, we rely on the sub-sample of cells for which we have information from two or more survey waves. Because a new sample of clusters is drawn for each DHS wave in the same country, it is only a subset of cells the country that is observed more than once. Table B.1 shows for both of our grid sizes the frequencies of countries, and of cells, that are observed in multiple waves.

Table B.1: Frequency of DHS waves for countries and cells

	(1)	(2)
Number of waves	Country frequency	Cell frequency
Panel A: Small grid cells		
2	11	11,512
3	12	3,249
4	1	488
5	1	297
Total	25	15,546
Panel A: PRIO-GRID cells		
2	11	2,406
3	12	1,153
4	1	146
5	1	39
Total	25	3,744

Notes: Table shows the number of countries (column (1)) and cells (column (2)) for which we have information from 2, 3, 4 or 5 geo-coded DHS waves in African countries in the years 1992–2013. Units of observation are 0.1×0.1 decimal degree grid cells in panel A, and 0.5×0.5 decimal degree PRIO-GRID cells in panel B.

The time intervals between two consecutive survey waves are typically three to eight years, but can be much larger for some countries. For example, geo-coded DHS for Togo are available for 1998 and for 2013, which makes a 15-years interval (the largest in our sample). For individual cells, intervals between two consecutive observations can be larger because a cell may be observed, for example, in the first and the last of three waves, but not in the second.

For cells that are observed more than once, we derive mean annual growth rates for each of our DHS-based development indicator ($r(DHS)$), for nighttime light ($r(light)$), and for electricity access ($r(electricity)$), assuming a constant annual growth rate between two survey waves. We use log differences in order to compute growth rates. In our main

specification, we add a small constant of 0.01 to all values before taking logs in order not to lose observations with value zero. For cells that are observed exactly twice, we have one mean annual growth rate. For cells that are observed three or more times, we calculate mean annual growth rates not only for intervals between consecutive survey waves, but for all intervals between each pair of survey waves. For example, for a cell that is observed in three different survey waves, we compute an average annual growth rate between the first and the second wave, between the second and the third wave, and between the first and the third wave.

This procedure assumes that each survey wave provides us with a good representation of the actual human development status within the cell area, so that we can derive local progress over time. However, consecutive DHS waves rarely sample the same village or urban block, let alone the same households. Our constructs of change in development over time may therefore suffer from substantial measurement error, because they may be confounded by time-invariant differences in development between localities sampled for different DHS waves, falling within the same cell. This concern becomes more important the larger our grid cells are: Within a 50×50 km zone, disparities in development are potentially higher than within a 10×10 km zone. In addition, there are differences across waves in terms of the number of clusters and the number of observations from which the cell average is computed. For example, a given cell may contain several survey clusters for one wave, but only a single cluster for the following wave. For simplicity, we treat the averages derived from the different bases as equivalent representations of the cell-level development status.

We include in this part of the analysis the same set of development indicators as for the cross-sectional analysis. Our empirical model regresses change in each of our development indicators on change in nighttime lights. Thereby, we include country fixed effects, allowing the relationship between change in development and change in nighttime lights to be country-specific. We also include fixed effects for the pair of years in which the base and end surveys were taken, to account for changes in satellite technology over the specific time period. We then introduce as an additional control the change in electricity access during the interval. In this model, while both the development indicators and electricity access are subject to measurement error due to different sample draws within the cell, change in nighttime lights is not. This is because we compute average nighttime lights within the full cell area for both the base and the end year. We cannot control for changes in population within the cell, given that we have no reliable measure for population growth at the cell level (see Section 2.3).

Table B.2 shows the results form our regressions of changes in development indicators on changes in nighttime lights within countries in a sample of all available survey pairs. We do not find any pattern which would suggest that an increase in nighttime lights (relative to the rest of the country) was associated with an increase in household wealth

or an improvement in education or health. The only exception is the positive association between changes in nighttime lights and changes in wealth when using the higher spatial resolution and not controlling for changes in electrification.

This non-result is robust to dropping all grid cells where our light variable is zero, to splitting the sample by the length of the time interval between any two surveys, or by dropping country fixed effects in order to investigate whether changes in nighttime lights are associated with changes in development across the African continent.

We conclude that while nighttime lights capture differences in human development in a cross-sectional comparison, changes in nighttime lights in a specific area over time are not associated with changes in DHS-based human development indicators in the same grid cells over the same time period. It is hard to interpret this absence of a measurable association. One interpretation could be that changes in nighttime lights over time are truly a bad proxy for changes in human development at the subnational level. A second interpretation could be that DHS are unsuitable to capture changes in local development over time as they are not designed as a panel, so that changes in our development indicators at the cell level are based on differences in development between households and survey clusters sampled for different DHS waves.

Table B.2: Changes in nighttime lights and changes in development within countries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep. var.:	r(wealth)		r(e-free wealth)		r(school attendance)		r(years of schooling)		r(infant mortality)		r(birth assistance)	
Panel A: Small grid cells												
r(light)	0.00821***	-0.00169	0.00169	-0.00383	-0.000746	-0.00669*	0.00927	-0.00695	0.0520	0.0645	-0.00794	-0.0241**
	(0.00244)	(0.00217)	(0.00266)	(0.00332)	(0.00417)	(0.00383)	(0.00593)	(0.00452)	(0.0676)	(0.0700)	(0.0120)	(0.00975)
r(electricity)		0.0893***		0.0487***		0.0596***		0.153***		-0.115*		0.156***
		(0.00392)		(0.0154)		(0.0117)		(0.0217)		(0.0661)		(0.0189)
Observations	16,758	16,758	16,328	16,328	20,220	20,096	20,245	20,117	20,502	20,502	19,194	19,194
R^2	0.051	0.281	0.065	0.100	0.165	0.177	0.054	0.119	0.071	0.071	0.072	0.107
Panel B: PRIO-GRID cells												
r(light)	0.0158	0.00363	0.00728	-0.00185	0.0194	-0.000765	0.00916	-0.0230	0.244	0.232	-0.0285	-0.0564**
	(0.0109)	(0.0113)	(0.0132)	(0.0143)	(0.0256)	(0.0271)	(0.0216)	(0.0213)	(0.192)	(0.200)	(0.0253)	(0.0241)
r(electricity)		0.0955***		0.0693***		0.135***		0.219***		0.0794		0.202***
		(0.00831)		(0.0149)		(0.0390)		(0.0293)		(0.0986)		(0.0266)
Observations	4,738	4,738	4,804	4,804	5,228	5,217	5,235	5,223	5,291	5,291	4,816	4,816
R^2	0.063	0.321	0.113	0.186	0.115	0.159	0.048	0.172	0.060	0.060	0.076	0.148

Notes: Linear regressions with country and year-pair fixed effects. Units of observation are cells of a 0.1×0.1 decimal degree spatial grid in panel A, and PRIO-GRID cells in panel B. All variables are mean annual growth rates between any two survey waves at which the same cell is observed, assuming a constant annual growth rate between the two survey waves. Growth rates are computed as log differences, after adding a small constant of 0.01 to the raw values. Standard errors are clustered at the country level. ***, **, * indicate significance at the 1, 5 and 10%-level, respectively.