UNU-WIDER | 30 YEARS OF RESEARCH FOR DEVELOPMENT

WIDER Working Paper 2015/085

Night lights and regional income inequality in Africa

Anthony Mveyange*

September 2015

Abstract: Estimating regional income inequality in Africa has been challenging due to the lack of reliable and consistent sub-national income data. I employ night lights data to circumvent this limitation. I find significant and positive associations between regional inequality visible through night lights and income in Africa. Thus, in the absence of income data, we can construct regional inequality proxies using night lights data. Further investigation on the night lights-based regional inequality trends reveals two main findings: first, increasing regional inequality trends between 1992 and 2003; and second, declining regional inequality trends between 2004 and 2012.

Keywords: regional income inequality, night lights, Africa
JEL classification: I132, R10, O550

Acknowledgements: I wish to thank UNU-WIDER for providing support and a stimulating environment during my PhD internship in Helsinki, Finland. I have enormously benefited from comments, critiques, and discussions with my PhD thesis supervisor Peter Sandholt Jensen, my UNU-WIDER mentor Channing Arndt, Jukka Pirttila, James Fenske, Adam Storeygard, Juan Miguel Villa, Smriti Sharma, Martina Kirchberger, Jennifer Kwok, Kuranda Morgan, and Cameron Breslin. I thank participants of the World Bank and Paris School of Economics 2014 Annual Bank Conference on Africa and University of Oxford's CSAE 2015 Annual Conference for their comments. I am grateful to participants during my preliminary job talk at SDU and SDU's PhD seminar series attendees for their constructive comments. Finally, special thanks to seminar participants at UNU-WIDER and Colorado State University for their insightful comments. An earlier version of this paper, circulated under the same name, used data from Brazil and South Africa to make a case for the night lights in estimating regional income inequality in Africa. This latest version replaces it. The usual caveats apply.

*Department of Business and Economics, University of Southern Denmark; amveyange@gmail.com.

This study has been prepared within the UNU-WIDER PhD internship programme.

ISSN 1798-7237 ISBN 978-92-9230-974-9 https://doi.org/10.35188/UNU-WIDER/2015/974-9

Typescript prepared by the author.

UNU-WIDER gratefully acknowledges the financial contributions to the research programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy-making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

UNU-WIDER, Katajanokanlaituri 6 B, 00160 Helsinki, Finland, wider.unu.edu

1 Introduction

For a long time the existing evidence on income inequality in Africa has been polarized between two extremes: at one extreme there is the micro[1] evidence that uses individuals and households as units of analysis and at the other extreme there is macro[2] evidence that uses countries and supra-nations as units of analysis. However, virtually no evidence exists on regional – the meso level – income inequalities in Africa as a whole. Kim (2008) asserts that the main reason for this vacuum has mainly been the lack of reliable and consistent sub-national income data.[3]

This paper offers an alternative approach that circumvents these data limitations to address the vacuum in regional income inequality (henceforth regional inequality) estimates in Africa. Thus, the paper has two main goals. The first goal seeks to exploit and show that night time satellite imagery data from the outer space are good proxies for approximating regional inequality when traditional sub-national income data are unavailable or unreliable, as is the case in Africa. The underlying hypothesis is that lights are good proxies for regional inequality in as much as they are proxies for income (Papaioannou, 2013; Sutton et al., 2007), economic growth (Henderson et al., 2012; Chen and Nordhaus, 2011), and wealth (Ebener et al., 2005). As detailed later, to show that lights are good proxies for estimating regional inequality, I base the analysis on Henderson et al. (2012)'s framework while restricting the geographical coverage to 423 regions across 32[4] countries in Africa for the years 1995, 2000, and 2005. The second goal aims at estimating, showing, and analysing the recent trends of regional inequality to gather insights on the underlying factors for the recent unbalanced regional growth distribution across countries in Africa (see for example WorldBank, 2012, for more details on this issue).

Several reasons motivate focusing the analysis at the regional level. First, regional inequality has been associated with conflicts and civil unrest in a number of African countries. However, studies in this area have had no success in measuring regional income inequality because of the lack of income data - they mainly relied on non-income and welfare measures of regional inequality. An excellent example of such studies is Østby et al. (2009) who used demographic data on household education and assets indicators to measure and decompose regional inequalities across African countries. Second, because regional inequality can also affect household income inequality (Kim, 2008), estimating it indirectly offers insights on the trends of household income inequality, which has been a subject of intense debates in Africa (see Deaton, 2005, for more details). Third, quantifying regional inequality trends also has clear policy implications especially when combined with growing concerns that the recent economic growth surges in Africa are not inclusive – the continent has witnessed a sharp increase in the number of poor people despite remarkable growth (WorldBank, 2012).

To achieve the stated goals, I confine the empirical analysis at the regional level (i.e., sec-

[1] See for example Tregenna and Tsela (2012); Deaton (2005); Chien and Ravallion (2001).

[2] See for instance Pinkovskiy and Sala-i Martin (2014a); Palma (2011); Sala-i Martin (2006).

[3] Not only that data at sub-national level are hard to come by, scholars even cast doubt on the quality of the existing data in Africa. For-example, Jerven (2014) asserts that national accounts statistics in Africa are problematic to the extent that they can be misleading in estimating economic activities. This concern is also shared by Deaton (2005), who documents serious problems with both national accounts and surveys data from the sub-Saharan Africa and considers them to be behind the recent empirical legacy of inconsistencies in poverty and inequality estimates in the region.

[4] See section 2.1 for an explanation on the choice of only 32 African countries.

ond level geographical administrative units) and divide it into two main parts. In the first part, I combine novel night lights (henceforth lights) data from the United States National Oceanic and Atmospheric Association's National Geophysical Data Center (NOAA-NGDC) available through its Defence Meteorological Satellite Program Operational Line-scan System (DMSP-OLS)[5] and income data from Tollefsen et al. (2012) derived from the original construction by Nordhaus et al. (2006). I use these two main datasets for (i) constructing measures of regional inequality across countries in Africa, and (ii) examining the relationship between lights-based and income based regional inequality indices[6] (i.e., Gini and Mean Logarithmic Deviations (MLD)). Applying panel fixed effects regressions, I find lights as good proxies for regional income disparities in the absence of income data. The findings suggest that a percentage point increase in light intensity Gini and MLD is associated with about 0.02 and 0.01 percentage points increase in income Gini and MLD respectively. These estimates are robust across a range of specification tests.

In the second part, building on the results above, I extend the analysis further by (i) estimating lights-based regional inequality indices in 748 regions[7] across 54 countries in Africa, and (ii) showing their trends across different geographical classifications during 1992–2012. I find four important patterns of regional inequality in Africa: (i) increasing regional inequality trends across regions in Africa during 1992–2003, (ii) declining regional inequality trends during 2004–2012, (iii) substantial variations across different regional groupings, indicating the sensitivity of inequality to regional and country differences, and (iv) the dominant role of the *between inequality* as a key driver of all these trends.

The present paper relates to a recent growing literature which uses lights data as a proxy for income, wealth, and economic growth in the absence of traditional income data. This literature is sub-divided into four strands[8]. First, Michalopoulos and Papaioannou (2013); Papaioannou (2013); Henderson et al. (2012) and Chen and Nordhaus (2011) use lights as a proxy for income and economic growth in sub-Saharan Africa. Second, Alesina et al. (2015) use lights to estimate ethnic inequality across countries. Third, Elvidge et al. (2012) develop a "night lights development index" to measure human development and track the distribution of wealth and income across countries. Fourth, Elvidge et al. (2009) and more recently Pinkovskiy and Sala-i Martin (2014b) employ lights to estimate poverty.

Relative to its predecessors, this paper contributes to this literature in several ways. First, it proposes that lights do a decent job in estimating regional disparities when income data are unavailable. Elvidge et al. (2012) offer a first attempt but find no meaningful association between lights and income-based inequality indicators. Their test is, however, at the country level with limited variation, and is based on time invariant cross-sectional set-up. As detailed later, the

[5] Available at: `http://ngdc.noaa.gov/eog/`

[6] The estimated indices are based on per capita, total and average income and light intensity, c.f. section 3.

[7] With lights data I am able to estimate more regions than the 423 regions sample used to validate that lights could be used as proxy for regional inequality c.f. section 3.

[8] Other studies that have documented the use of lights to approximate economic activities at the sub-national level include Hodler and Raschky (2014) who use data on the intensity of lights as a proxy for economic activities and, hence, GDP growth across 126 countries to estimate regional favoritism. For other parts of the world, Villa (2014) uses lights to approximate the growth of Colombian municipalities, Levin and Duke (2012) compare Israel and the West Bank to show that differences in lights reflect the underlying differences in subnational socio-economic activities across the two countries.

analysis in this paper, on the contrary, not only covers regions but also exploits panel gridded data to estimate regional inequality over time. Second, this paper shows the average trends of regional inequality across countries for a relatively longer period (i.e., between 1992 and 2012), a new addition to the existing literature. Third, Elvidge et al. (2009) and Pinkovskiy and Sala-i Martin (2014b) have recently focused on estimating poverty and not regional inequality, as is the case in this paper. Fourth, the paper is consistent with Alesina et al. (2015) in that the lights data are used to construct inequality measures. However, it is different for three main reasons: (i) it measures regional inequality not ethnic inequality, (ii) it explores variation in regional inequality at local scales over a relatively long period (two decades) compared to Alesina et al. (2015)'s focus at country level over a relatively short period, (iii) it checks the validity of lights-based inequality indices not addressed by Alesina et al. (2015). Fifth, this paper includes lights-based decomposable measures of regional inequality to identify the sources of the observed regional inequality, an element absent in the existing literature.

Section 2 describes the conceptual framework between lights and regional inequality. This section also demonstrates how regional inequality indices are calculated. Section 3 describes the main data used for the empirical analysis. Section 4 presents the empirical framework and econometric estimation. Section 5 presents the main baseline regression estimates, followed by Section 6 which presents the sensitivity checks to the baseline regressions. Section 7 presents the estimates of lights-based regional inequality indices. Section 8 concludes.

2 Lights and Regional Inequality

The conceptual framework of this paper builds on two important assertions by Henderson et al. (2012):

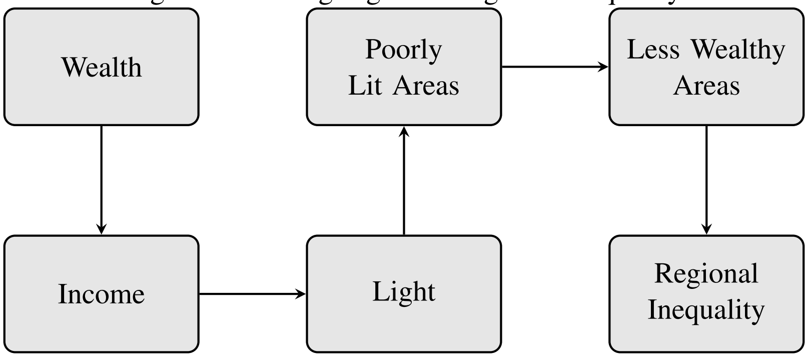
> *Intensity of night lights reflects outdoor and some indoor use of lights. More generally, however, consumption of nearly all goods in the evening requires lights. As income rises, so does lights usage per person, in both consumption activities and many investment activities. Obviously, this is a complex relationship, and we abstract from such issues as public versus private lighting, relative contributions of consumption versus investment, and the relationship between daytime and night time consumption and investment...(p.999).*

and

> *...lights in an area reflect total intensity of income, which is increasing in both income per person and number of people...(p.1001)*

These assertions inform the conceptual framework of the paper, which is summarized in Figure 1. The figure conveys the underlying theoretical conjecture that relates to Elvidge et al. (2009, p.1653), who distinctly offer an interesting implicit insight to the relationship between lights and regional inequality. That is, they assert that "areas with higher population counts in developing countries would be poorly lit and therefore have higher percentages of poor people (lights being considered as a proxy for wealth)." The most direct implication of this assumption, of interest in this paper, is that to the extent that lights are positive and strong correlates of income, poorly lit regions tend to have low income and are less wealthy. Therefore, similar to Elvidge et al. (2011) but different in scope, this paper partly investigates the extent to which this relationship is both true and meaningful. In other words, as previously noted, the paper examines the extent to which lights-based regional inequality indicators are good proxies for income-based regional inequality indicators.

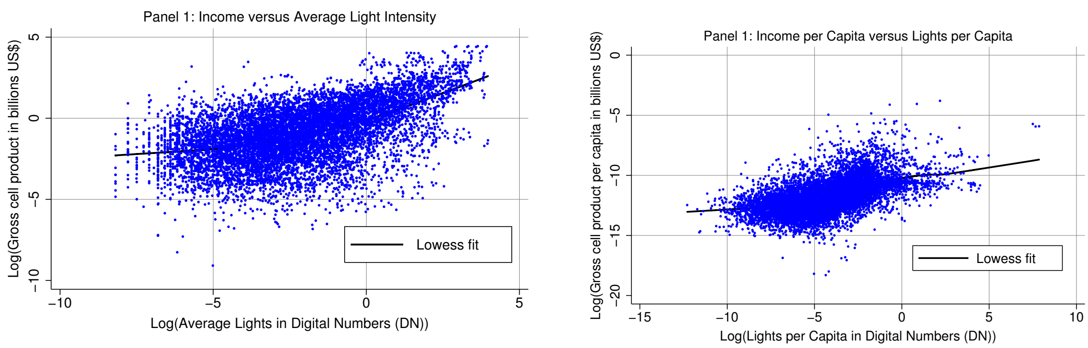
Figure 1: Linking Lights to Regional Inequality



Source: Author's illustration.

The arrows in Figure 1 illustrate the connection between wealth, income, lights, and regional inequality, assuming other background factors remain unchanged. That is, wealth predicts the levels of income which in turn determines the intensity of lights. This relationship implies that relative to areas with high light intensity, poorly lit areas will tend to be less wealthy. This second relationship offers a measure of *regional inequality*.[9]

Figure 2: Scatter Plots of Grid Cell Lights and Income for the Years 1995, 2000, and 2005.



Source: Author's calculations.

The conceptual framework suggests that the relationship between income and light intensity is monotonic. However, it prompts two important questions: (i) whether the relationship between a measure of income and a measure of light intensity is linear, (ii) whether the relationship between a measure of income inequality and a measure of light intensity inequality is linear. Answers to these questions are of first order importance. Similar to Henderson et al. (2012) I investigate the first question by employing both lights and G-Econ data at 0.5 decimal degree

[9]In general, this simple conceptual framework can apply for an analysis at different geographical administrative units, for example, grid cells, counties, districts, or regions.

grid cells across countries in Africa. Figure 2 depicts a lowess fit of lights (measured as average light intensity) and total income (measured as total gross cell product). The figure shows a somewhat linear relationship[10] between average light intensity and income. A similar lowess fit on the second question also shows a linear relationship[11] c.f. figures 3 and 4 in section 5.

Lights are, however, not a perfect measure of economic activity. As noted by Henderson et al. (2012); Elvidge et al. (2011) and Chen and Nordhaus (2011) saturation, over-glow and blooming are potential problems.

Saturation occurs primarily in developed countries in which the intensity of lights is high and the inherent top-coding of lights data is problematic because data censoring limits inference beyond the earmarked thresholds. Over-glowing occurs because as lights travel from one point to another its reflection can wrongly be recorded as originating from particular area. Blooming occurs primarily in places where the likelihood of observing completely dark places is high - e.g., in poor countries. Believing that these practical limitations can bias the inference of these data is not unreasonable, particularly when extended for regional inequality analysis.

In their recent study, however, Michalopoulos and Papaioannou (2013) show that saturation and over-glowing are trivial across African countries making them less of a threat. Blooming, however, remains a potential threat especially in Africa, where the likelihood of observing dark places is high. In the empirical model I account for this by controlling for unlit grid cells within a region. The next subsection applies the deduced conceptual framework to show the calculations of the regional inequality indices.

2.1 Regional Inequality Indices

As noted previously, to calculate regional inequality indices, I employ Tollefsen et al. (2012)'s gridded spatial data which are terrestrial 0.5×0.5 decimal degree[12] grid cells estimates of output and population across countries in the world. These spatially rescaled data are derived from 1×1 decimal degrees G-Econ[13] data originally constructed by Nordhaus et al. (2006). I spatially join these gridded data with level 2 geographical administrative units (normally referred to as provinces or regions) data extracted from the global administrative areas (GADM)[14] covering 32 countries in Africa. I refer to these dataset as gridded regional income data. Tollefsen et al. (2012) also offer polygon feature with 0.5×0.5 grid cells. This polygon feature is used to extract lights and spatially join with regions across the 32 countries. I then pair up these gridded regional lights data with the gridded regional income data. The final compiled data is used to calculate regional inequality indices across countries[15] for the years 1995, 2000, and 2005, c.f.

[10] The unreported raw correlation coefficient is 0.52 between income and average light intensity and 0.57 between income per capita and lights per capita intensity.

[11] This finding is unsurprising: the consumption of lights (e.g., generated from electricity) is an expenditure that requires income to sustain. Therefore, it is not unreasonable to observe a linear relationship underlying the distribution of inequality in income and light intensity c.f. Section 4.

[12] This is equivalent to 55×55 square kilometres, or 3025 area size which is about the size of a county or district. In this sense, the grid cells are treated as second level geographical administrative units.

[13] Available at: `http://gecon.yale.edu/`

[14] Available at: `http://www.gadm.org/`

[15] In the early versions of the paper I explored the variation of lights and income data at districts or counties level in a specific region. However, this requires adjusting the area sizes in calculating regional inequality indices. The choice of grid cells, though considered relatively arbitrary and noisy, is consistent in terms of the area size uniformity across cells in respective regions across countries making it more consistent and less prone to measurement errors.

Section 3.

Figure 12 in the appendices offers insights behind the regional inequality indices calculations. The figure shows the grid cells nested within regions across 32 African countries. Country borders are in maroon and regional borders are in blue. The maroon bordered white empty countries are the 22 countries that are excluded in the analysis. These are countries whose regions are either smaller than 0.5 × 0.5 decimal degree (e.g., countries like Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, Mauritius, Comoros, and Seychelles) or have few grids that limit meaningful variation (e.g., Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia, and Eritrea). The analysis also excludes Somalia – a war ravaged country in Africa – despite being large enough for analysis. As noted above, in each region the grid cell's spatial and temporal variation in both lights and income data is used in calculating regional income inequality over time.

To estimate regional inequality, I calculate two standard measures of inequality, Gini and Mean Logarithmic Deviations (MLD), for both income and lights data. Gini is commonly used as a standard measure of income inequality across individuals, households, regions, and countries. Similarly, MLD is used for the same purpose but measures the dispersion of income at the lower tail of income distribution which, in this case, overlaps with Africa's low within regions distribution of income. Also, the choice of MLD is appealing because it allows inequality to be decomposed into *within* and *between* components, which are important elements in understanding the variation of total income inequality across subgroups such as households, regions, or countries (Anand, 1983). Equations 1 and 2 show the calculations of these indices for each region across countries over the years 1995, 2000, and 2005.

$$Gini = \frac{\sum_{i=1}^{n}(2i-n-1)*y_i}{n^2P} \tag{1}$$

$$MLD^{16} = \frac{1}{N}*\sum ln\left(\frac{\bar{y}_i}{y_i}\right) \tag{2}$$

where i is grid cell rank order, n total number of grid cells, y_i is grid cell value (i.e., per capita and total lights or per capita and total income), $\bar{y}_i$ is the average grid cell lights or income value, and P is grid cell population average.

Equation 3 decomposes MLD into its two main components: the *within* and the *between* inequality. In the context of regions, the within component captures the variation of income inequality within regions mainly because of grid level variations in income and light intensity. The between component, on the contrary, is that part of total income inequality that explains the variations between regions.

$$MLD^{17} = \sum V_k * MLD_k + \sum V_k * ln\left(\frac{\bar{y}}{\bar{y}_k}\right) \tag{3}$$

where V_k subgroup k population share, $\bar{y}$ is overall mean lights or income value, and $\bar{y}_k$ is the

[16] The formula comes from Haughton and Khandker (2009).

[17] Also derived from Haughton and Khandker (2009).

mean lights or income value in sub-group k. As explained later, these subgroups include sub-continental division of Africa, coastal versus landlocked countries, and mineral rich and – poor countries. Others include countries favourable to agriculture versus those that are not, and countries classified based on their previous colonial regime. MLD_k is inequality for subgroup k calculated as if the subgroup were a separate population. Thus, the first element on the right hand side captures the within inequality component. The second component captures the between inequality which is derived assuming every grid within a given subgroup k has k's mean income or light intensity.

3 Data

3.1 Night Lights Data

Lights data are extracted, using ArcGIS software, from the United States Defense Meteorological Satellite Program Operational Line-scan System (DMSP-OLS)[18]. These data are reported as 30 arc-second cells which are equivalent to 0.86 square kilometres at the equator. The data are recorded by satellites orbiting the earth every day between 20:00 and 22:30 local time across countries and are available since 1992. These data are in two main formats: (i) the average visible and stable lights free from cloud coverage and (ii) the average lights with the percentage frequency of lights detection[19]. Most of the economic applications of lights data use the former format, as is the case in this paper.

The visible and stable lights data are further classified into three types. The first is the cloud free lights imagery data that, as Lowe (2014) asserts, are useful for identifying areas with low numbers of observations where the data quality is demeaned. The second contains raw lights data which have not been filtered for auroral or ephemeral events and other background noises. The third contains stable lights data that have been cleaned up of all auroral or ephemeral events and background noises. I use the stable lights data, which are recorded in digital numbers from 0 (no lights) to 63 (high lights intensity), for the calculation of lights Gini and MLD. I follow Lowe (2014) for extracting and cleaning these data including masking out the geographical areas with observed gas flares.

3.2 Income Data

As mentioned above, Income data come from Tollefsen et al. (2012) who spatially rescale Nordhaus et al. (2006)'s 1×1 decimal degree spatial output data – commonly known as G-Econ data – into 0.5×0.5 decimal degrees grid cells. The G-Econ estimates gross product (i.e., gross domestic product) and population data across countries available in 5 years interval between 1990 and 2005. For Africa and other poor countries, the estimation of population counts at lower geographical administrative units rely on national census data. For the estimation of gross cell product, the dataset relies on both national income and non-income data. For example, in some countries the estimation is based on sectoral (e.g., agricultural and non-agricultural) employment data. In other cases involving resource rich countries the analysis is extended to include oil and

[18]Available at: http://ngdc.noaa.gov/eog/

[19]More details are available at: http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

mining production (especially if the production was more than 10 percent of country's GDP) (see Nordhaus et al., 2006, pp. 10 – 14).

However, a key limitation with the G-Econ dataset is that the construction of income data relies on very strong assumptions; for-example when non-income data are used to estimate GDP of grid cells across country. To reinforce this point, Nordhaus et al. (2006, p. 12) asserts that "because the resolution of the economic data is so poor, we judge these estimates to be relatively unreliable". Yet, G-Econ dataset is somewhat useful in analysing regional income and inequality dynamics in Africa. Besides, this dataset is the only available source offering comprehensive gridded income data covering all countries in the continent.

The most important variables in this dataset are the total gross cell product (henceforth total income) and gross cell product per capita (hence forth income per capita) calculated as gross cell product (total gross cell product in a grid) divided by total population in a grid cell. These variables are thus equivalent to total income and income per capita in a given grid cell available for the years 1990[20], 1995, 2000 and 2005.

3.3 Population Data

I use two main sources for extracting population data. First, the gridded population of the world (GPW) data, available in the G-Econ data, converted into 0.5 decimal degree grid cells by Tollefsen et al. (2012). Similar to income, I also pair these population count data with regions across countries for years 1990, 1995, 2000, and 2005. To combine these data with yearly lights data, I follow Hodler and Raschky (2014) and linearly interpolate and extrapolate the available regional gridded population counts over the missing years. Finally, the compiled data contains yearly population count estimates between 1992 and 2005.

To get gridded population counts from 2006 to 2012, I resort to the Landscan[21] spatial global dataset[22] which is a 30 arc-second (a 0.86 square kilometre resolution) grid product containing global population counts between 2000 and 2012. Combining these data with the interpolated population counts above, I end up with regional gridded population counts dataset covering the period between 1992 and 2012.

3.4 Other Data

I use two more data sources. The first is the global administrative areas database (GADM)[23]. I employ this data source to extract polygon features for all levels 2 (i.e., regions or provinces) geographical administrative units across countries. I then use it to individually pair gridded lights, income and population counts data before combining all of them together. To ensure that I address potential grids-region overlapping I invoke ArcGIS spatial join feature "completely contains" which matches grids within regions in the lights dataset if they are completely contained in the regions of the gridded income and population counts data.

[20] Since lights data are available from 1992, I exclude this year in the empirical analysis.

[21] Note that I deliberately use data from 2006 to avoid the concerns that previous versions of Landscan data used light intensity in calibrating population counts.

[22] Archived by Oak Ridge National Laboratory: `http://web.ornl.gov/sci/landscan/landscan_data_avail.shtml`

[23] Available at: `http://www.gadm.org/`

The second source is the Brazilian statistical bureau[24] which archives GDP[25] and population data for over 5000 municipalities across 26 states in Brazil. The analysis with these data is restricted for the 2000–2010 periods in which all the data are consistently available. I mainly use these data for the robustness checks of the correlations between income and lights-based regional inequality indices. Note that the estimates of regional inequality indices (both for income and lights) are reported at the state level. As detailed in Section 6, the calculations use the variation of lights and income at municipality level to estimate state level inequality indices.

3.5 Summary Statistics

Table 1 reports the summary statistics. The variances of the indices show reasonable variability of the indices around their mean. However, inequality indices based on G-Econ data are, on average, significantly lower than those based on lights data. Of course, one cannot expect indices from the two sources to be the same. There are two possible explanation for this difference. One possible explanation can be aligned with Nordhaus et al. (2006)'s assertion that G-Econ data for Africa are somewhat unreliable and the data generating process relies on very strong assumptions. Assuming that data construction is not an issue, a second plausible reason for this difference, as put by Pinkovskiy and Sala-i Martin (2014b), is the inherent different data generating processes behind both G-econ and lights data. Summaries on lights data show significant variability of lights variables around their mean – the standard deviation is twice the mean for all variables. Further, the summaries suggest that blooming appears to be a potential problem: there are, on average, about 14 unlit grid cells in a region for a given year.

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Gini - total gross cell product	1235	0.246	0.158	0.0	0.795
Gini - sum of light intensity	1235	0.514	0.302	0.0	0.996
Gini - gross cell product per capita	1235	0.080	0.109	0.0	0.593
Gini - light intensity per capita	1235	0.508	0.311	0.0	0.996
Gini - mean light intensity	1235	0.522	0.301	0.0	0.996
MLD - total gross cell product	1235	0.225	0.266	0.0	1.914
MLD - sum of light intensity	1235	0.516	0.517	0.0	2.955
MLD - gross cell product per capita	1235	0.046	0.118	0.0	1.503
MLD - light intensity per capita	1235	0.479	0.577	0.0	4.629
MLD - mean light intensity	1235	0.550	0.554	0.0	2.955
Light intensity per capita (DN)	1235	0.131	1.695	0.0	36.929
Sum of light intensity (DN)	1235	22737	48964	0.0	467301
Mean light intensity (DN)	1235	2.042	5.613	0.0	49.897
Unlit grid cells	1235	13.581	27.637	0.0	205

Note: The averages are for years 1995, 2000 and 2005 in 423 regions across 32 African countries. DN stands for the digital number, which are censored between 0 and 63.

[24] Available at: `http://www.ipeadata.gov.br/`
[25] Reported at constant 2000 prices

4 Empirical Strategy

In this section, I present the empirical model estimating the relationship between income and lights-based regional inequality indicators. As alluded to before, the model is inspired by Henderson et al. (2012, pp. 1005–1009). Since data are available for the years 1995, 2000, and 2005 I run panel fixed effect regressions using an unbalanced panel of 423 regions across 32 African countries in the sample. The identification is based on within regions variations in regional inequality indices. The estimated baseline models are given as:

$$Gini_{i,j,t}^{Income} = \gamma_1 Gini_{i,j,t}^{Lights} + X'_{i,j,t}\gamma_2 + \Gamma_j + \Psi_t + \varepsilon_{i,j,t}, \tag{4}$$

$$MLD_{i,j,t}^{Income} = \tau'_1 MLD_{i,j,t}^{Lights} + X'_{i,j,t}\tau_2 + \theta_j + \zeta_t + \nu_{i,j,t}, \tag{5}$$

where i is country, j is region, and t is years 1995, 2000, and 2005. The main outcome variables for equations 4 and 5 are based on G-Econ data and calculated using both income per capita and total income variables. Similarly, the covariates are based on lights data calculated using both lights per capita and sum of lights intensity.

The vector **X** controls for observed confounding factors. Two potential factors can confound the estimates. The first is the unlit grid cells in regions across countries and over time. Unlit grid cells reflects the previously described blooming problem and have the potential to confound the coefficient lights estimates upward. I include the unlit grid cells in the regression to thwart this potential upward bias. Second, for the lit grid cells the spatial and temporal variation in light intensity within a region also have the potential to bias the coefficient estimates: places with dispersed light intensity are likely to have more inequality relative to places with less dispersion in lights intensity. As argued before, I assume a linear relationship between lights and income-based regional inequality indices. The validity of this assumption necessitates accounting for the dispersion in light intensity. The baseline analysis thus controls for the standard deviation of light intensity across regions over the years 1995, 2000, and 2005.

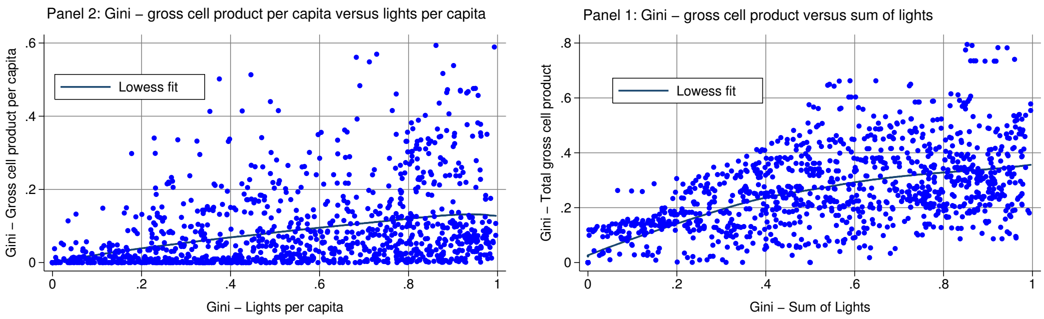
The empirical model also accounts for the unobserved regional fixed effects across countries. As suggested by Henderson et al. (2012), these include such factors as climate, regional economic activities and conditions, cultural factors that influence use of lights, public and private lightning, and electricity generating conditions. Regional fixed effects are represented by Γ_j and θ_j in the models. The empirical specification also accounts for the time fixed effects to account for time varying influences such as changes in income, consumption of lights, and patterns of regional economic activities over time. Time fixed effects are represented by Ψ_t and ζ_t in the models. ε and ν are error terms. Finally, the model controls for intra-cluster correlation and heteroskedasticity by clustering the the standard errors of coefficient estimates at the regional level.

5 Results

5.1 Correlations: Income versus Lights Inequality

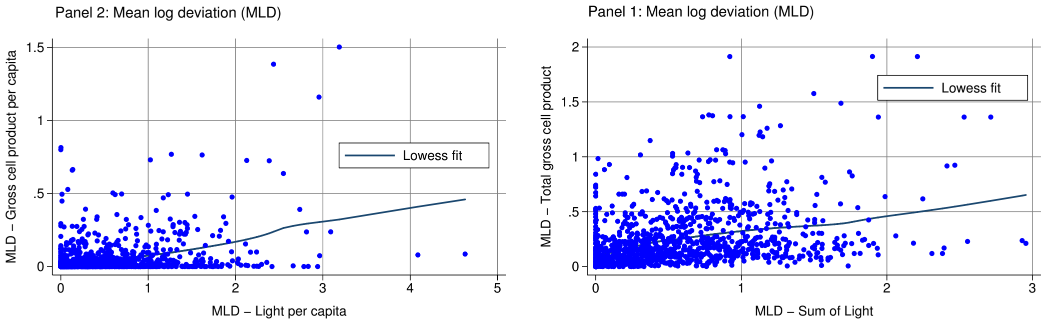
To show lights as good proxies for estimating regional inequality, Figure 3 presents two panels showing the correlation scatter plots between per capita income Gini and per capita lights Gini; and total income Gini and Gini based on the sum of lights. Figure 4 shows the two similar panels but with MLD correlations. As the lowess fit on the data indicates, the main observation to emerge from these figures suggest a positive correlation of the inequality indicators.

Figure 3: Gini Lowess Scatter Plots



Source: Author's calculations

Figure 4: MLD Lowess Scatter Plots



Source: Author's calculations

Mapping these correlations, Figures 8 and 9 in the appendices present a visual spatial distributions of average Gini based on G-Econ and lights data across 432 regions in 32 countries in Africa during 1995–2005. Gini in Figure 8 is based on both per capita income and light intensity whereas in Figure 9 it is based on both total income and light intensity. Similarly, Figures 10 and 11 (c.f. the appendices) present a visual spatial distributions of average MLD based on G-Econ and lights data during 1995–2005. As above, MLD in Figure 10 is based on both per capita income and light intensity while that in Figure 11 is based on both total income and light intensity. A general observation from all these figures shows that countries with high lights-based regional inequality tend to also have high income-based regional inequality. The converse holds true.

Nonetheless, this visual inspection is silent about the magnitude of the correlation coefficients which are relevant for gauging whether lights-based regional inequality indicators have predictive power over the income-based indicators. Table 11 (c.f. the appendices) reports the correlation coefficients between lights and income-based Gini and MLD. The correlation between per capita income and lights Gini is 0.4 while the correlation between Gini based on total income and sum of light intensity is 0.6. Similarly, the correlation between per capita income and lights MLD is 0.3 while the correlation between MLD based on total income and sum of light intensity is 0.4. Moreover, the table also reports the correlations of regional inequality indicators based on average light intensity relative to those based on total and per capita income. In general, the correlations of both Gini and MLD closely remain unchanged relative to above correlations.

In summary, these results indicate statistically significant and reasonable associations between income and lights-based regional inequality indices. As noted before, the first goal of this paper is to establish the relationship between lights and regional inequality indices and to draw conclusions on whether the observed association is meaningful. The next subsection presents the results from the baseline regression estimates.

5.2 Predicting Regional Inequality with Lights

Table 2 reports the baseline regression estimates based on per capita income Gini as an outcome variable. The table shows the regression estimates with lights per capita (columns 1 and 3); sum of lights (columns 2 and 5); and average lights (columns 3 and 6) as predictors. Columns 1 – 3 report the estimates with the inclusion of lights standard deviation as the only control. Columns 4 – 6 (which show the main baseline results) control for both lights standard deviation and the number of unlit grid cells.

Table 2: Income Per Capita Gini versus Light Gini

	Gini	Gini	Gini	Gini	Gini	Gini
Lights Gini - DN per capita	0.018^{*}			0.021^{**}		
	(0.009)			(0.010)		
Lights Gini - total DN		0.023^{**}			0.026^{**}	
		(0.011)			(0.011)	
Lights Gini - average DN			0.023^{**}			0.026^{**}
			(0.011)			(0.011)
Lights Standard Deviation	0.002^{***}	0.002^{***}	0.002^{***}	0.002^{**}	0.002^{**}	0.002^{**}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.009	0.011	0.011	0.013	0.015	0.015
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is income per capita Gini from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000 and 2005. (v) The unlit cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

The regression estimates indicate that a percentage point increase in lights per capita Gini increases per capita income Gini by 0.021 percentage points. The results also show a small increase to 0.026 increase in per capita income Gini for a point increase in both sum and average lights Gini. As shown in Table 12 in the appendices, these effects are equivalent to 0.049, 0.038, and 0.042 standard deviations increase in per capita income Gini for a standard deviation increase in per capita, sum and average lights Gini, respectively.

Although the magnitude of the effect is fairly moderate, the estimates are statistically significant supporting the claim that regional inequality in light intensity predicts the regional inequality in income per capita. The results also indicate that the dispersion of light intensity around its mean is statistically significant and has a positive relationship – of a small magnitude – with per capita income Gini across all specifications. This result is unsurprising: if lights proxy income, high light intensity variances will be associated with significant disparities in regional inequality.

Table 3: Total Income Gini versus Lights Gini

	Gini	Gini	Gini	Gini	Gini	Gini
Lights Gini - total DN	0.018*			0.020**		
	(0.009)			(0.010)		
Lights Gini - DN per capita		0.023**			0.025**	
		(0.010)			(0.010)	
Lights Gini - average DN			0.020**			0.022**
			(0.010)			(0.010)
Lights Standard Deviation	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.012	0.014	0.013	0.014	0.017	0.015
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is total income Gini from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (v) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principe, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 3 reports the estimates based on total income Gini as outcome variables. The columns are ordered as in Table 2. The results remain unchanged with and without the inclusion of the relevant controls. Moreover, the standardized estimates, in Table 13 in the appendices, show 0.038, 0.049, and 0.042 standard deviations increase in total income Gini for a standard deviation increase in Gini based on sum, per capita and average lights Gini. Again, the results corroborate the above findings: light intensity disparities are good proxies for estimating regional income disparities in the absence of income data.

Table 4 presents the estimates based on per capita income and lights MLD. The columns are organized as in Table 2. In general, the estimates indicate statistically significant and positive predictive power of lights-based per capita MLD on per-capita income MLD. The sizes, fairly small in magnitude, of the effects are 0.010 percentage point increase in income per capita MLD for a point increase per capita lights MLD. The results also show a small decrease to 0.007 and 0.008 in per capita income MLD for a point increase in both sum and average lights MLD, respectively. These point estimates are equivalent to 0.051, 0.032, and 0.039 standard deviations increase in per capita income MLD for a standard deviation increase in per capita, sum and average lights MLD respectively (c.f. Table 14 in the appendices). The effects of the variance of light intensity is also similar to baseline Gini estimates. Overall, this table also supports the claim that lights are good proxy for predicting regional inequality when income data are absent.

Table 4: Income Per Capita MLD versus Lights MLD

	MLD	MLD	MLD	MLD	MLD	MLD
Lights MLD - DN per capita	0.011***			0.010**		
	(0.004)			(0.004)		
Lights MLD - total DN		0.008**			0.007**	
		(0.003)			(0.003)	
Lights MLD - average DN			0.009***			0.008***
			(0.003)			(0.003)
Lights Standard Deviation	0.003**	0.003**	0.003**	0.003**	0.003**	0.003**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.046	0.031	0.035	0.047	0.032	0.036
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is income per capita MLD from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (v) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 5 reports the estimates based on total income MLD. Again, the columns are organized as above. The fairly moderate coefficient size estimates indicate statistically significant and positive prediction of MLD based on the sum of lights and average lights on total income MLD. The results on lights per capita MLD though positive, are marginally significant. The results show that the variability of light intensity around its mean strongly and positively affects the total income MLD. Further, as shown in Table 15 in the appendices, the standardized estimates are roughly 0.02 standard deviations increase in total income MLD for a standard deviation increase in MLD based of sum, per capita and average light intensity. The overall picture presented by these results is also in line with the claim proposed above.

To summarize this section, the regression estimates show light intensity inequality significantly and positively predicts regional income inequality. Furthermore, the estimates show a moderate magnitude of the prediction between light intensity and regional inequality.

Table 5: Total Income MLD versus Lights MLD

	MLD	MLD	MLD	MLD	MLD	MLD
Lights MLD - total DN	0.010**			0.010**		
	(0.005)			(0.005)		
Lights MLD - DN per capita		0.009*			0.009*	
		(0.005)			(0.005)	
Lights MLD - average DN			0.010**			0.010**
			(0.004)			(0.004)
Lights Standard Deviation	0.002**	0.002**	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.016	0.018	0.017	0.016	0.018	0.017
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is gross cell product (i.e., total income) MLD from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. The outcome variable is gross cell product (i.e., total income Gini). (v) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

6 Robustness Checks

This section demonstrates the qualitative and quantitative robustness of the baseline results. I test the robustness of baseline estimates in four distinct ways. First, I check the robustness of the correlations presented above. Second, I refit the baseline model using different data to show a different set of estimates supporting the baseline findings. Third, I tweak the baseline model specification to check the sensitiveness of the coefficient estimates. In the last step, I refit the baseline model using a full sample of countries assuming that all countries are large enough for inequality estimation at 0.5×0.5 grid cells. The first two ways use data from Brazil whereas the last two use the baseline data.

As noted in section 3.4, I combine the Brazilian municipality GDP and lights data to test both correlations and regression estimates for Gini and MLD. The units of analysis in this case are 26 Brazilian States which are each equivalent to level 2 geographical administrative units[26]. I therefore construct a dataset with lights and income-based inequality indices for the period 2000-2010 across 26 Brazilian States. Since Brazil is in the upper middle income category, I also calculate and include in the regressions the *Theil T index* (henceforth Theil) which is useful for analysing income and lights at the upper-end of of their distributions. Obviously,

[26]Note that I calculated State inequality indices differently: instead of using grid cells variations, I used over 5000 municipalities' GDP and light intensity variations to estimate State level Gini and MLD. I chose this method to ensure comparability of regional inequality indices from both lights and GDP data at the State level since the GDP data are based on political jurisdictions. As highlighted previously, the main limitation of this approach, relative to grid cells analysis, is non-uniformity of the municipality areas. Yet, it suffices for checking the robustness of the baseline results.

given Brazilian income status, MLD may fail to fully account for the dynamics of state level inequality in the country.

Table 6: GDP versus Lights State Inequality Correlations - Brazil

	$Gini_{GDP}$	$Gini_{Lights}$	MLD_{GDP}	MLD_{Lights}	$Theil_{GDP}$	$Theil_{Lights}$
$Gini_{GDP}$	1					
$Gini_{Lights}$	0.744^{***}	1				
MLD_{GDP}	0.706^{***}	0.557^{***}	1			
MLD_{Lights}	0.257^{***}	0.727^{***}	0.480^{***}	1		
$Theil_{GDP}$	0.941^{***}	0.729^{***}	0.848^{***}	0.390^{***}	1	
$Theil_{Lights}$	0.538^{***}	0.934^{***}	0.508^{***}	0.889^{***}	0.606^{***}	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6 presents the correlation of these indices. The coefficient between income and lights Gini is 0.744. Similarly, the coefficient for MLD and Theil index are 0.480 and 0.606 respectively. These correlations are significant and show strong associations between income and lights-based inequality indicators. Therefore, they are robust to the baseline correlations estimates, signifying that the relationship between income and lights-based inequality indices is not only positive but also shows that lights are good proxies for regional inequality indices.

Table 7: Robustness checks - Income versus Lights State Inequality in Brazil

	Income Gini	Income MLD	Income Theil
Lights - Gini	0.038^{*}		
	(0.022)		
Lights - MLD		0.035	
		(0.029)	
Lights - Theil			0.053^{*}
			(0.030)
Lights Standard Deviation	0.002	-0.001	0.003
	(0.002)	(0.011)	(0.007)
Unlit Grid Cells Controls	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Observations	286	286	286
R-squared	0.104	0.165	0.170
States	26	26	26
Regional Fixed Effects	Yes	Yes	Yes

Notes: (i) The outcome variables are Gini, MLD and Theil index from based on Brazilian GDP data. (ii) Standard errors are clustered at the state level. (iii) The regressions are based on panel data between 2000 and 2010. (iv) The unlit cells refer to the total municipalities with zero light intensity in digital numbers (DN). * $p < .10$, ** $p < .05$, *** $p < .01$

Table 7 shows the regressions estimates after refitting the baseline models using data from Brazil. The results show that a point increase in lights-based state inequality indicators (i.e., Gini, MLD and Theil) increases Income Gini, MLD and Theil by 0.038, 0.035, and 0.053 respectively. However, these point estimates are marginally significant for Gini and Theil and insignificant for MLD. The marginal significance of Gini and Theil may partly be due to small sample size. The insignificance of MLD is unsurprising: Brazil is an upper middle income country so MLD may not capture the full dynamics of state level inequality. Gini and Theil in this case are realistic indicators to look at in describing the relationship between income and lights inequality at the state level. Overall, although the Gini and Theil estimates are marginally significant they still convey the same message as the baseline estimates: lights are good proxies of regional inequality in the absence of sub-national income data.

Table 8: Robustness Checks - Income Per Capita Gini & MLD versus Lights Gini & MLD

	Gini	Gini	Gini	MLD	MLD	MLD
Lights Gini - DN per capita	0.020^{**}					
	(0.010)					
Lights Gini - total DN		0.025^{**}				
		(0.011)				
Lights Gini - average DN			0.025^{**}			
			(0.011)			
Lights MLD - DN per capita				0.011^{**}		
				(0.004)		
Lights MLD - total DN					0.008^{***}	
					(0.003)	
Lights MLD - average DN						0.009^{***}
						(0.003)
Lights - DN per capita	0.006^{**}			0.006		
	(0.003)			(0.004)		
Lights - total DN		$2.27e-07^{***}$			$2.83e-07^{***}$	
		$(6.09e-08)$			$(8.97e-08)$	
Lights - average DN			0.001^{**}			0.001
			$(4.04e-04)$			(0.001)
Unlit Grid Cells Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.013	0.026	0.014	0.042	0.038	0.030
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) All the outcome variables (i.e., Gini and MLD) are based on income per capita from G-Econ. (iii) Standard errors are clustered at the regional level. (iv) *DN* stands for digital numbers. (v) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (vi) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vii) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Other countries include Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea excluded for having fewer grid cells which limits meaningful variation. (viii) I also excluded Somalia because of its persistent conflicts thus making the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

I now turn to refitting the baseline regression models by tweaking the control variables: instead of using standard deviation of lights as control I use the variation in the sum of lights, lights per capita, and the average lights. The idea is to test whether the results are sensitive to the changes of control variables.

Tables 8 reports the estimates based on per capita income and lights Gini and MLD. Columns 1 – 3 show the estimates based on Gini. Columns 4 – 6 show the estimates based on MLD. Overall, the results show strong and positive relationship between lights and income-based indices, with the magnitude of the prediction close to that in the baseline estimates.

Table 9: Robustness Checks - Total Income Gini & MLD versus Lights Gini & MLD

	Gini	Gini	Gini	MLD	MLD	MLD
Lights Gini - total DN	0.019^{**}					
	(0.009)					
Lights Gini - DN per capita		0.024^{**}				
		(0.010)				
Lights Gini - average DN			0.021^{**}			
			(0.010)			
Lights MLD - total DN				0.010^{**}		
				(0.005)		
Lights MLD - DN per capita					0.009^{*}	
					(0.005)	
Lights MLD - average DN						0.010^{**}
						(0.004)
Lights - total DN	$1.05e-07^{**}$			$1.96e-07^{**}$		
	$(5.04e-08)$			$(9.22e-08)$		
Lights - DN per capita		0.002			0.003	
		(0.002)			(0.004)	
Lights - average DN			$3.34e-04$			$3.06e-04$
			$(2.87e-04)$			(0.001)
Unlit Grid Cells Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.013	0.012	0.011	0.018	0.017	0.016
F-Statistics [p-value]	0.057	0.080	0.086	0.040	0.335	0.126
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) All the outcome variables (i.e., Gini and MLD) are based on total income from G-Econ. (iii) Standard errors are clustered at the regional level. (iv) *DN* stands for digital numbers. (v) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (vi) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vii) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Other countries include Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea excluded for having fewer grid cells which limits meaningful variation. (viii) I also excluded Somalia because of its persistent conflicts thus making the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Tables 9 reports the estimates based on total income and lights Gini and MLD. Columns 1 – 3 show the estimates based on Gini. Columns 4 – 6 show the estimates based on MLD. Similar to Table 8 the results also show significant and positive relationship between lights and income-based indices, with the size of the prediction fairly close to the baseline estimates.

Finally, I re-estimate the baseline regressions using the full sample of 51 African countries[27]. With the exception of the relationship between sum of lights Gini and total income Gini, the general picture from the remaining results support a claim that lights data are good proxies for estimating regional income inequality in the absence of income data. Tables 17, 19, 16, and 18

[27]The analysis excluded Sao Tome and Principe, Eritrea and Mauritius because of poor data quality.

in the appendices show these results.

In summary, all the robustness results show reasonable positive associations, in some cases with closely similar magnitudes, between income and lights-based inequality indices. These results corroborate the estimates and conclusions of baseline regressions.

Up to this point, I have exploited and shown that night lights data do a decent job as proxies for regional income inequality in the absence of traditional income data in 432 regions across the 32 countries sample in Africa. As stated in the introduction this was the first goal of this paper. In the following section, building on the above findings, I embark on the second objective: to estimate, show and analyse the trends of regional inequality in Africa from 1992 to 2012.

7 Estimating Regional Income Inequality with Lights Data

Similar to Pinkovskiy and Sala-i Martin (2014a), Table 10 shows the distribution of regions across different classifications for all African countries. These regions are used for estimating income inequality for the period 1992–2012. In total, the estimation and hence the analysis takes into account 748 regions[28] across all 54 African countries.

Table 10: Regions by Geographical Subdivisions

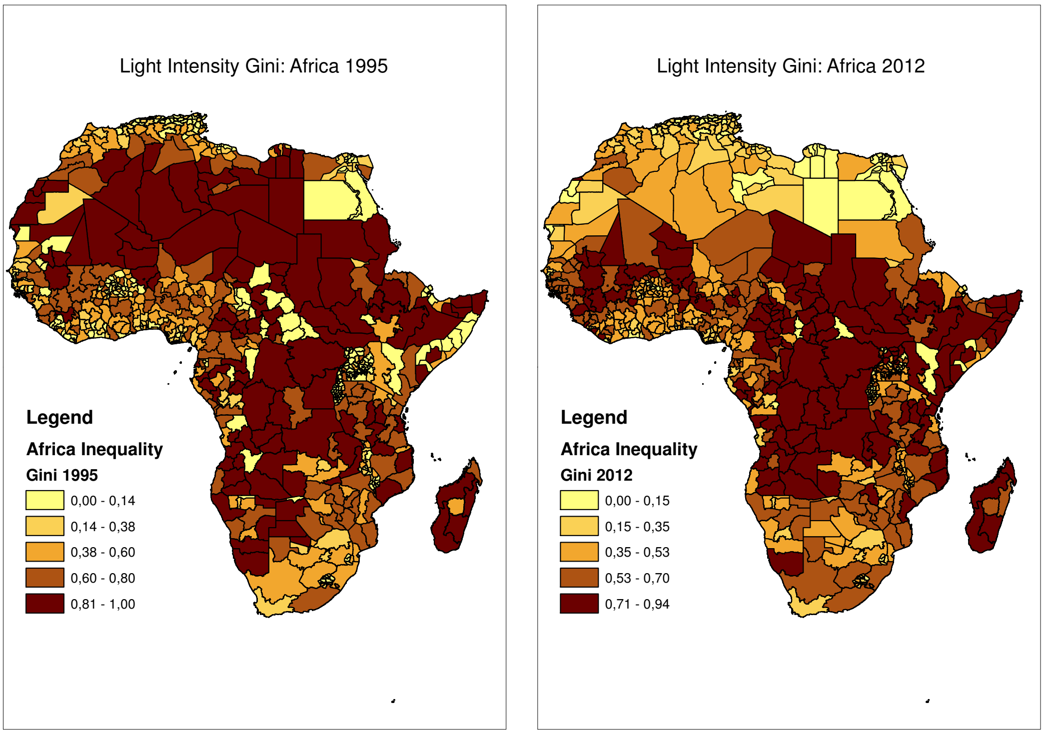
Category	Classification	Number of regions	Number of countries
Sub continent	*Eastern*	223	17
	Central	100	9
	Northern	152	7
	Southern	45	5
	Western	228	16
Geography	*Coastal*	518	40
	Landlocked	230	14
Mineral	*Rich*	355	23
	Poor	393	29
Agriculture	*Favourable*	295	25
	Unfavourable	453	27
Colonial master	*British*	433	27
	French	223	17
	Portugal	56	5
	Belgium	36	3
Total		**748**	**54**

Note: This analysis includes also Sao Tome and Principe, South Sudan, and Eritrea, thus 54 countries.

Proceeding with the estimation, Figure 5 shows a snapshot of the distribution of regional inequality across regions in African countries in the sample in 1995 and 2012. The figure suggests that most regions in the northern Africa had less regional inequality 17 years later in 2012 than in 1995: the spatial Gini is within the first to the third quantile (i.e. 0.00 to 0.53) in 2012 compared to the fourth and firth quantiles (i.e., 0.60 to 1.00) in 1995. A somewhat similar pattern holds

[28] For limitations noted previously, the baseline regression estimates are based on a restricted sample, – i.e., 423 regions. Lights data, however, do not suffer from these limitations thus allowing a coverage of large regional sample.

Figure 5: Regional Inequality in 1995 and 2012 in Africa



Source: Author's calculations

for some regions in the western and south-western Africa.

There are, however, some regions that experienced no change or modest changes in regional inequality between these two periods. Some regions in the central and eastern Africa experienced virtually no change (Gini quantiles in 1995 were closely similar to those in 2012) while others experienced modest increases, from the fourth in 1995 to the fifth in 2012, Gini quantiles. Regions in the southern Africa also experienced both modest declines and increases during these two periods, with the most variations moving back and forth between the third, fourth and fifth Gini quantiles. The overall picture, however, shows that spatial Gini in the third, fourth, and fifth quantiles are relatively higher in 1995 than in 2012, suggesting a modest decline in regional inequality in Africa in these two periods.

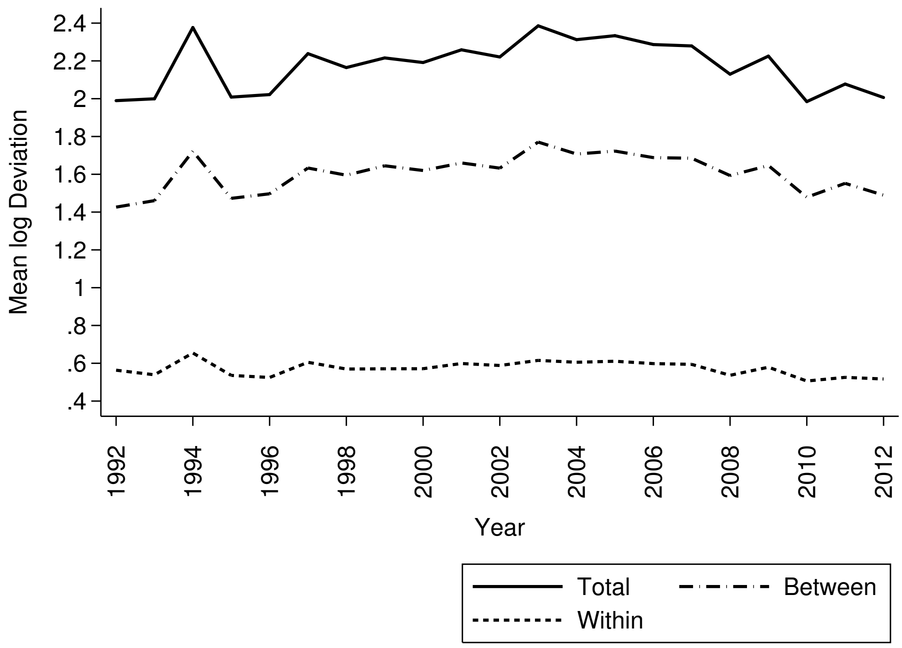
To track the average trends over time across regions and countries, Figure 6 shows the average lights-based Gini and MLD for the period between 1992 and 2012 in the 748 regions across the 54 African countries. Visual inspection of the figure suggests two main phenomena at work in the last two decades in Africa. First, in the first decade (i.e., between 1992 and 2003) regional inequality as measured by Gini coefficient in panel 1 increased close to 0.78 Gini points. A sharp decline observed around 1994 and 1995 perhaps suggests the initial effect of economic growth as documented by Pinkovskiy and Sala-i Martin (2014a), who argue that growth spurts in Africa started around this time. However, a persistent increase afterwards can thus be associated with the subsequent first-order detrimental effects of economic growth on regional inequality in the continent.

Second, post-2003 the trends started to decline and reached 0.72 Gini points in 2012. This mild reduction can be understood to relate to the second-order effects of economic growth in

Figure 6: Gini Trends in Africa; 1992–2012

Source: Author's calculations

Figure 7: MLD Trends in Africa; 1992–2012



Source: Author's calculations

Africa in which economic growth imposes regional convergences in inequality perhaps because of spillovers and agglomeration effects. Similar to the first decade, this decade also had some declines; with a sharp one between 2005 and 2006 and a slight upswing during 2008–2009, which was the peak of the second wave of the recent financial crisis. Arguably, it is not unreasonable to associate economic growth as a primary predictor of regional inequality in Africa. Moreover, other potential factors could also play a similar role. For-example, social, geographical and climatic, cultural and other unobserved factors could be potential explanations for the observed regional inequality trends.

The regional inequality trends in total MLD as shown in Figure 7 also reveal somewhat similar trends as Gini, with the decline being driven primarily by between region inequality suggesting the importance of regional differences in explaining regional inequality in Africa. The within region inequality is unchanged, an indication that the contribution of within region inequality to overall inequality is rather low. However, these trends vary when I further subdivide African countries into different categories c.f. Figures 14 - 27 in the appendices.

Figures 13, 14, 15, 16, and 17 in the appendices show these trends when African countries are grouped based on subcontinental groupings. Overall, Panel 1 indicates that regional inequality (as measured by Gini) trends have generally been on the decline in the period 1992–2012, with slight variations specific to regional groupings. For example, except for central Africa (where inequality rose between 1992 and 2003 and then started falling from 2004 onwards) and eastern Africa (where oscillatory movements are observed in the period 1992–2003 and steady declines afterwards), regional income inequality has generally been on the decline for the rest of the regional groupings.

Decomposing these regional inequality trends, Panel 2 indicates a dominant role of between inequality as captured by MLD. The most striking observation from these trends is that whereas central and eastern experienced a sharp increase at the peak of the second wave of the recent financial crisis in 2008 and 2009, northern and southern African countries experienced no change in their regional inequality levels and western African countries experienced a slight decline.

Figures 18 and 19 present the trends of Gini and MLD in coastal and landlocked countries. Comparing these trends reveals that Gini was high and virtually remained unchanged in the period 1992–2003 before starting to fall steadily afterwards for both coastal and landlocked countries. However, countries in both categories experienced a sharp increase between 2008 and 2009, followed by a decline between 2009 and 2010, perhaps suggesting the subsiding effects of the financial crisis. Similar to geographical groupings, the MLD suggests that most of the variations in total inequality were driven by inequality between regions.

Comparisons of regional inequality trends between mineral poor and mineral rich countries show consistent patterns. Figures 20 and 21 summarize these trends. On the one hand, Figure 20 indicates the unchanged average Gini patterns for mineral rich countries until after 2003 when it started to fall steadily. On the other hand, Figure 21 reports a slight increase in Gini for the period 1992–2005 followed by sharp declines in 2006 and 2009 then a relatively stable pattern afterwards. As with other groupings discussed above, between region inequality dominates in explaining the evolution of total regional inequality in this category as well.

Turning to countries that are favourable to agriculture and those that are not, Figures 22 and 23 summarize the regional inequality patterns. Whereas, Figure 22 indicates a general decline in regional inequality since 1992, Figure 23 shows a general increase in Gini between 1992 and 2005 followed by a sharp decline in 2006. Overall, all the figures have one thing in common: Gini rose during 2008-2009 and sharply declined during 2009-2010, with between region inequality being the main predictor of total MLD.

To this point, we have said nothing about the regional inequality trends based on countries' colonial origin. Figure 24 presents the trends for British colonies' countries, which suggest a modest average rise in regional inequality between 1992 and 2005, followed by a sharp decline in 2006, with movements in the total MLD being dominated by between region inequality. Figure 25 shows trends in the former French colonies. The picture is somewhat similar to former British colonies: Gini rose until 2006 and then sharply declined. Similar to British colonies, total MLD in former French colonies was driven by between region inequality.

Finally, Figures 26 and 27 summarize the inequality trends for former Portuguese and Belgian colonies, respectively. Overall, the trends of Gini appear to indicate a persistent increase, again with between region inequality being the main driver of total MLD. However, unlike the Belgium colonies and all other groupings, within region inequality appears to also explain total MLD in the Portuguese colonies.

In general, five main conclusions can be drawn from these trends. First, mixed evidence on regional inequality trends between 1992 and 2003, with some country groupings experiencing declines and others increases in inequality. Second, declining patterns in regional inequality was for most country groupings recorded post-2003. When compared with the recent economic growth spurts in Africa, these two conclusions are thus consistent with the famous Kuznets (1955)'s hypothesis. Third, the mixed patterns of inequality, mostly between 2008 and 2009 when the second wave of the financial crisis was at its peak, and 2009–2010 period when the effects started to subside. Fourth, the trends suggests stable movements in regional inequality after 2010. Fifth, the trends unambiguously show that regional inequality is primarily dominated by between region rather than within region inequality. This last point suggests two important policy lessons. First, within region inequality across African countries has been rather low for the past two decades. Second, between region inequality has been the dominating force in explaining the evolution of regional inequality in Africa during the same period.

8 Conclusions

This paper explores whether night lights data are useful in estimating regional inequality in Africa where income data are persistently unavailable or unreliable. Building on their recent use as proxies for income, wealth, and growth, the paper shows strong, positive, and robust relationship between lights and income-based regional inequality indicators. These results support the claim that night lights are good proxies in estimating regional inequality in the absence of tradition income data, as is the case for Africa.

Proceeding to estimate and analyse the regional inequality trends using night lights data, the present paper also offers the first systematic empirical estimation and analysis of regional inequality trends in Africa over the past two decades. The analysis shows that regional inequality was on the rise in the first decade between 1992 and 2003 and started to fall steadily between 2004 and 2012 across countries in Africa. Further investigation reveals the role of between region inequality as a key driver of these trends. Moreover, the findings show variations in regional inequality trends across different classifications, indicating the sensitivity of inequality to country and regional differences. Overall, the analysis signals the importance of incorporating the dimension of spatial differences in policy discussions as a tool for spurring and spreading balanced spatial economic growth across countries in Africa.

To summarize, this paper shows that lights data are potential proxies for estimating regional inequality in Africa where lack of income data is persistent. However, I do not claim that lights data fully capture the income inequality dynamics in Africa. Obviously, the data still have their own practical limitations, a somewhat different data-generating process, and may be associated with strong assumptions for their use. Yet, working with these data while cautiously observing their building blocks enables this paper to set a broader context for policy and further research. This is not only relevant in Africa and but also in other developing regions where sub-national income data are unavailable and unreliable and where regional inequality has recently become a concern and focus for sound economic policy. Specifically, the use of lights data to estimate regional inequality offers two possibilities for future research: (i) a possibility of comparing both regional and household inequality in a way that assesses the extent to which the former contributes to the latter in Africa. This area is potentially under-explored and calls for further research, and (ii) now that it is possible to estimate regional inequality trends over a two-decades period, further empirical work could help in understanding the main determinants of such trends within and between regions across countries in Africa.

References

Alesina, A. F., S. Michalopoulos, and E. Papaioannou (2015, June). Ethnic inequality. *Journal of Political Economy 123*(3), 547 – 724.

Anand, S. (1983). *Inequality and poverty in Malaysia: Measurement and decomposition*. World Bank.

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

Chien, S. and M. Ravallion (2001). How did the world's poorest fare in the 1990s? *Review of Income and Wealth 47*(3), 283–300.

Deaton, A. (2005). Measuring poverty in a growing world (or measuring growth in a poor world). *Review of Economics and Statistics 87*(1), 1–19.

Ebener, S., C. Murray, A. Tandon, and C. C. Elvidge (2005). From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery. *International Journal of Health Geographics 4*(1), 5.

Elvidge, C., K. Baugh, S. Anderson, P. Sutton, and T. Ghosh (2012). The night light development index (nldi): a spatially explicit measure of human development from satellite data. *Social Geography 7*(1).

Elvidge, C., P. Sutton, K. Baugh, D. Ziskin, T. Ghosh, and S. Anderson (2011). National trends in satellite observed lighting: 1992-2009. In *AGU Fall Meeting Abstracts*, Volume 1, pp. 03.

Elvidge, C. D., P. C. Sutton, T. Ghosh, B. T. Tuttle, K. E. Baugh, B. Bhaduri, and E. Bright (2009). A global poverty map derived from satellite data. *Computers & Geosciences 35*(8), 1652–1660.

Haughton, J. H. and S. R. Khandker (2009). *Handbook on poverty and inequality*. World Bank Publications.

Henderson, J. V., A. Storeygard, and D. N. Weil (2012). Measuring economic growth from outer space. *American Economic Review 102*(2), 994–1028.

Hodler, R. and P. A. Raschky (2014). Regional favoritism*. *The Quarterly Journal of Economics 129 (2)*, 995–1033.

Jerven, M. (September, 2014). African growth miracle or statistical tragedy? interpreting trends in the data over the past two decades. *UNU-WIDER, WIDER Working Paper 2014/114*.

Kim, S. (2008). Spatial inequality and economic development: Theories, facts, and policies. *Urbanization and Growth*, 133–166.

Kuznets, S. (1955). Economic growth and income inequality. *The American economic review*, 1–28.

Levin, N. and Y. Duke (2012). High spatial resolution night-time light images for demographic and socio-economic studies. *Remote Sensing of Environment 119*, 1–10.

Lowe, M. (2014). Night lights and arcgis: A brief guide. Technical report, MIT.

Michalopoulos, S. and E. Papaioannou (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica 81*(1), 113–152.

Nordhaus, W., Q. Azam, D. Corderi, K. Hood, N. M. Victor, M. Mohammed, A. Miltner, and J. Weiss (2006). The g-econ database on gridded output: methods and data. *Unpublished manuscript, Yale University 12*.

Østby, G., R. Nordås, and J. K. Rød (2009). Regional inequalities and civil conflict in sub-saharan africa1. *International Studies Quarterly 53*(2), 301–324.

Palma, J. G. (2011). Homogeneous middles vs. heterogeneous tails, and the end of the 'inverted-u': It's all about the share of the rich. *Development and Change 42*(1), 87–153.

Papaioannou, E. (2013). National institutions and subnational development in africa. *The Quarterly Journal of Economics 129*(1), 151–213.

Pinkovskiy, M. and X. Sala-i Martin (2014a, September). Africa is on time. *Journal of Economic Growth 19*(3), 311–338.

Pinkovskiy, M. and X. Sala-i Martin (2014b). Lights, camera,... income!: Estimating poverty using national accounts, survey means, and lights. Technical report, National Bureau of Economic Research.

Sala-i Martin, X. (2006). The world distribution of income: falling poverty andĚ convergence, period. *The Quarterly Journal of Economics 121*(2), 351–397.

Sutton, P. C., C. D. 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*(S07), 5–21.

Tollefsen, A. F., H. Strand, and H. Buhaug (2012). Prio-grid: A unified spatial data structure. *Journal of Peace Research 49*(2), 363–374.

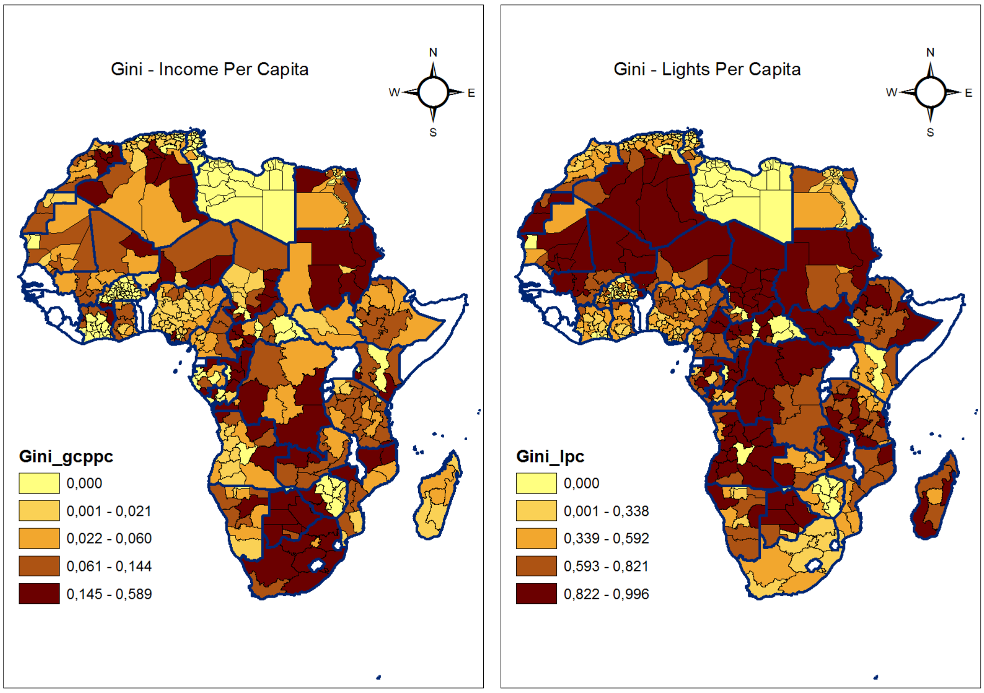
Tregenna, F. and M. Tsela (2012). Inequality in south africa: The distribution of income, expenditure and earnings. *Development Southern Africa 29*(1), 35–61.

Villa, J. M. (2014). Social transfer and growth: The missing evidence from luminosity. Technical report, UNU-WIDER.

WorldBank (2012). *World Development Indicators 2012*. World Bank Publications.

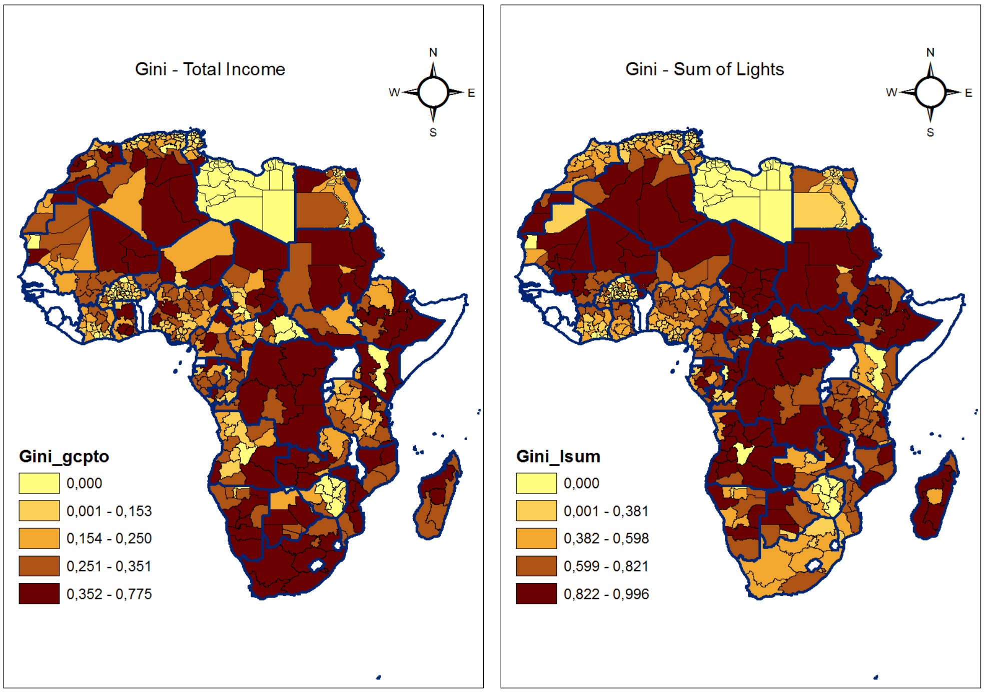
Appendices

Figure 8: Average Income Per Capita and Lights Per Capita Gini: 1995–2005



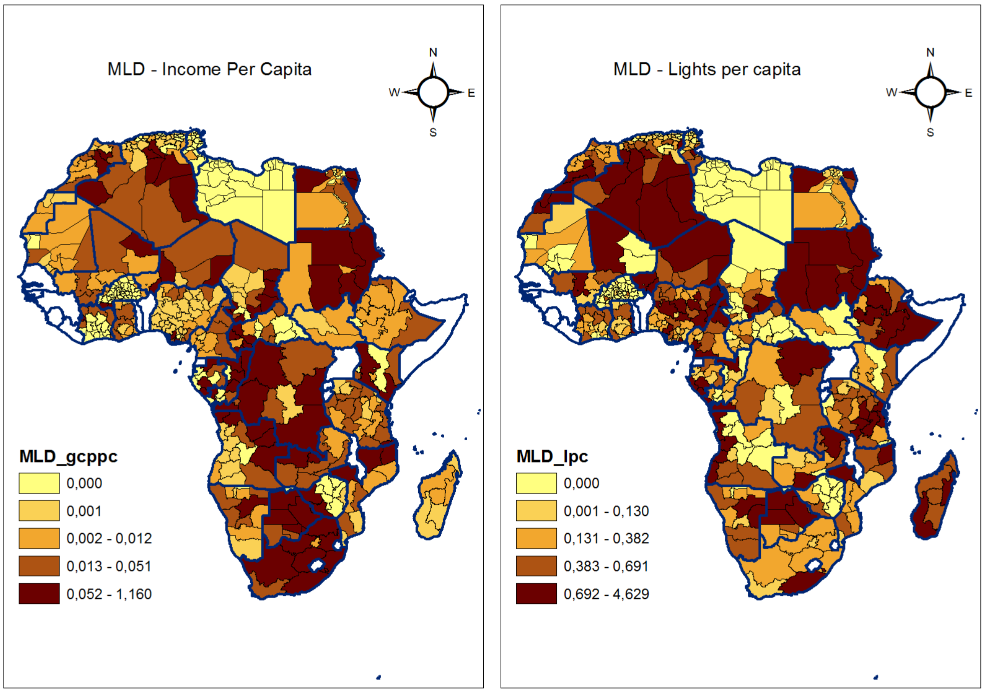
Source: Author's calculations.

Figure 9: Average Total Income and Sum of Lights Gini: 1995 -2005



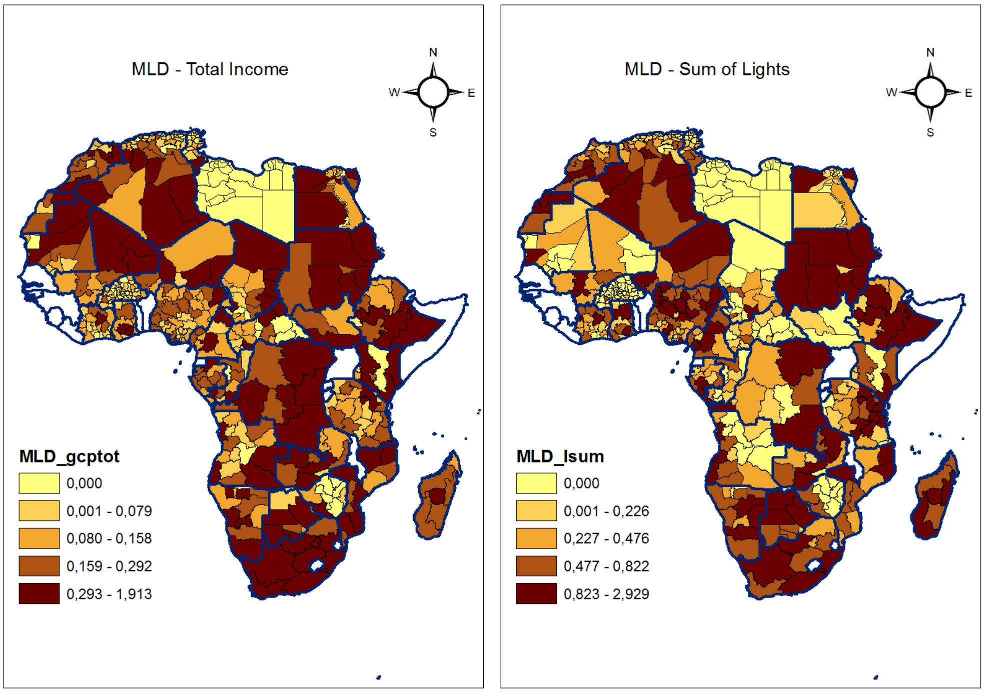
Source: Author's calculations.

Figure 10: Average Income Per Capita and Lights Per Capita MLD: 1995–2005



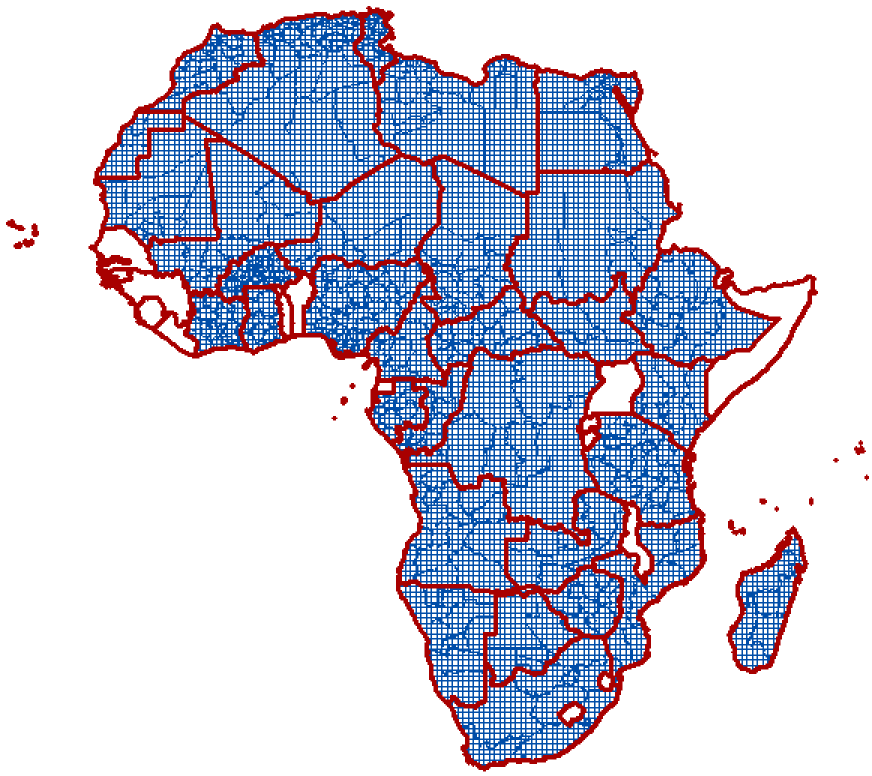
Source: Author’s calculations.

Figure 11: Average Total Income and Sum of Lights MLD: 1995 -2005



Source: Author’s calculations.

Figure 12: Grid Cells and Regions in Africa



Source: Author's calculations.

Table 11: Correlation Table: G-Econ versus Lights Data

	Gini_gcppc	Gini_lpc	Gini_gcptot	Gini_lsum	Gini_lmean	MLD_gcppc	MLD_lpc	MLD_gcptot	MLD_lsum	MLD_lmean
Gini_gcppc	1									
Gini_lpc	0.392***	1								
Gini_gcptot	0.563***	0.569***	1							
Gini_lsum	0.396***	0.969***	0.621***	1						
Gini_lmean	0.397***	0.965***	0.623***	0.987***	1					
MLD_gcppc	0.843***	0.270***	0.374***	0.265***	0.267***	1				
MLD_lpc	0.323***	0.591***	0.418***	0.555***	0.555***	0.319***	1			
MLD_gcptot	0.504***	0.273***	0.844***	0.326***	0.325***	0.354***	0.256***	1		
MLD_lsum	0.311***	0.491***	0.526***	0.546***	0.539***	0.255***	0.843***	0.386***	1	
MLD_lmean	0.309***	0.480***	0.522***	0.524***	0.551***	0.260***	0.824***	0.380***	0.958***	1

gcppc is gross cell product per capita; *gcptot* is total gross cell product; *lsum* is sum of lights intensity; *lmean* is mean lights intensity; and *lpc* is lights intensity per capita

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Income Per Capita Gini versus Light Gini - Standardized Estimates

	Gini	Gini	Gini	Gini	Gini	Gini
Lights Gini - DN per capita	0.045**			0.049**		
	(0.019)			(0.019)		
Lights Gini - total DN		0.034*			0.038**	
		(0.018)			(0.018)	
Lights Gini - average DN			0.039**			0.042**
			(0.019)			(0.019)
Light Standard Deviation	0.011***	0.011***	0.011***	0.011***	0.011***	0.010***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.014	0.012	0.013	0.017	0.014	0.015
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is income per capita Gini from G-Econ data. (ii) All the variables are standardized with mean 0 and standard deviation 1. (iii) Standard errors are clustered at the regional level. (iv) *DN* stands for digital numbers. (v) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000 and 2005. (vi) The unlit cells refer to the total grid cells with zero light intensity digital numbers (DN). (vii) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 13: Total Income Gini versus Lights Gini - Standardized Estimates

	Gini	Gini	Gini	Gini	Gini	Gini
Lights Gini - total DN	0.034*			0.038**		
	(0.018)			(0.018)		
Lights Gini - DN per capita		0.045**			0.049**	
		(0.019)			(0.019)	
Lights Gini - average DN			0.039**			0.042**
			(0.019)			(0.019)
Lights Standard Deviation	0.011***	0.011***	0.011***	0.011***	0.011***	0.010***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.012	0.014	0.013	0.014	0.017	0.015
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is total income Gini from G-Econ data. (ii) All the variables are standardized with mean 0 and standard deviation 1. (iii) Standard errors are clustered at the regional level. (iv) *DN* stands for digital numbers. (v) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (vi) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vii) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 14: Income Per Capita MLD versus Lights MLD - Standardized Estimates

	MLD	MLD	MLD	MLD	MLD	MLD
Lights MLD - DN per capita	0.053***			0.051**		
	(0.020)			(0.020)		
Lights MLD - total DN		0.036**			0.032**	
		(0.015)			(0.013)	
Lights MLD - average DN			0.043***			0.039***
			(0.016)			(0.014)
Lights Standard Deviation	0.027**	0.028**	0.028**	0.026**	0.027**	0.027**
	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.046	0.031	0.035	0.047	0.032	0.036
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is income per capita MLD from G-Econ data. (ii) All the variables are standardized with mean 0 and standard deviation 1. (iii) Standard errors are clustered at the regional level. (iv) *DN* stands for digital numbers. (v) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (vi) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vii) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 15: Total Income MLD versus Lights MLD - Standardized Estimates

	MLD	MLD	MLD	MLD	MLD	MLD
Lights MLD - total DN	0.019**			0.019**		
	(0.009)			(0.009)		
Lights MLD - DN per capita		0.019*			0.019*	
		(0.011)			(0.011)	
Lights MLD - average DN			0.020**			0.020**
			(0.009)			(0.009)
Lights Standard Deviation	0.007**	0.007**	0.007**	0.007**	0.007**	0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1235	1235	1235	1235	1235	1235
R-squared	0.016	0.018	0.017	0.016	0.018	0.017
Countries	32	32	32	32	32	32
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is gross cell product (i.e., total income) MLD from G-Econ data. (ii) All the variables are standardized with mean 0 and standard deviation 1. (iii) Standard errors are clustered at the regional level. (iv) *DN* stands for digital numbers. (v) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. The outcome variable is gross cell product (i.e., total income Gini). (vi) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vii) Countries excluded in the analysis are Togo, Benin, Gambia, Malawi, Guinea Bissau, Djibouti, Equatorial Guinea, Sao Tome and Principle, Rwanda, Lesotho, Swaziland, Burundi, and islands such as Mauritius, Comoros and Seychelles. Senegal, Guinea, Sierra Leone, Liberia, Uganda, Tunisia and Eritrea are excluded for having few grid cells, which limits meaningful variation. (vii) I also excluded Somalia because of its persistent conflicts, which make the reliability of data highly questionable. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 16: Full Sample - Income Per Capita Gini versus Light Gini

	Gini	Gini	Gini	Gini	Gini	Gini
Lights Gini - DN per capita	0.025***			0.026***		
	(0.009)			(0.009)		
Lights Gini - total DN		0.028***			0.030***	
		(0.010)			(0.010)	
Lights Gini - average DN			0.027***			0.028***
			(0.010)			(0.010)
Lights Standard Deviation	0.002***	0.002***	0.001**	0.001**	0.002***	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1636	1636	1636	1636	1636	1636
R-squared	0.013	0.015	0.014	0.014	0.016	0.015
Countries	51	51	51	51	51	51
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is gross cell product per capita (i.e., income per capita) MLD from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (v) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) This analysis includes 51 African Countries. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 17: Full Sample - Total Income Gini versus Lights Gini

	Gini	Gini	Gini	Gini	Gini	Gini
Lights Gini - total DN	0.012			0.013		
	(0.009)			(0.009)		
Lights Gini - DN per capita		0.018**			0.020**	
		(0.009)			(0.009)	
Lights Gini - average DN			0.015*			0.017*
			(0.009)			(0.009)
Lights Standard Deviation	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1636	1636	1636	1636	1636	1636
R-squared	0.008	0.010	0.009	0.009	0.011	0.010
Countries	51	51	51	51	51	51
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is gross cell product (i.e., total income) Gini from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (v) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) This analysis includes 51 African Countries. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 18: Full Sample - Income Per Capita MLD versus Lights MLD

	MLD	MLD	MLD	MLD	MLD	MLD
Lights MLD - DN per capita	0.012***			0.011***		
	(0.004)			(0.004)		
Lights MLD - total DN		0.010**			0.009**	
		(0.004)			(0.005)	
Lights MLD - average DN			0.011**			0.011**
			(0.005)			(0.005)
Lights Standard Deviation	0.003**	0.003**	0.003**	0.003**	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1636	1636	1636	1636	1636	1636
R-squared	0.044	0.032	0.038	0.044	0.033	0.038
Countries	51	51	51	51	51	51
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

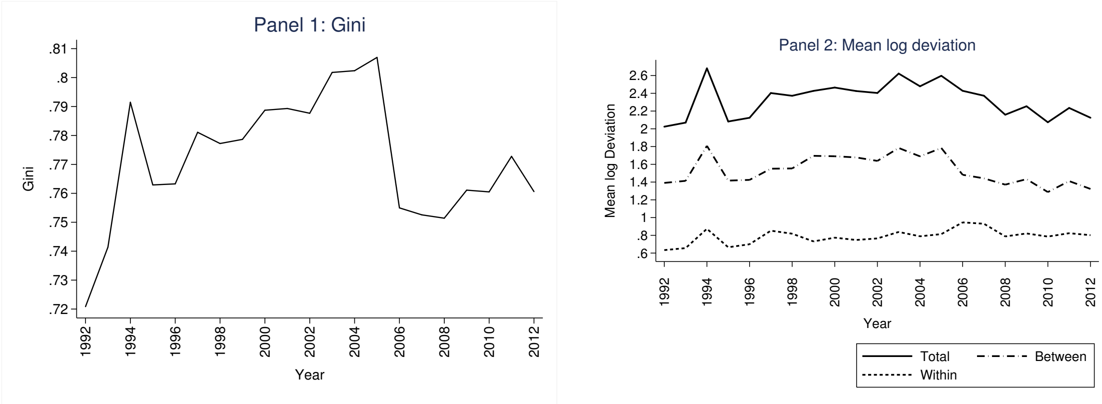
Notes: (i) The outcome variable is gross cell product per capita (i.e., income per capita) MLD from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (v) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) This analysis includes 51 African Countries. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 19: Full Sample - Total Income MLD versus Lights MLD

	MLD	MLD	MLD	MLD	MLD	MLD
Lights MLD - total DN	0.015**			0.015**		
	(0.006)			(0.007)		
Lights MLD - DN per capita		0.016**			0.016**	
		(0.007)			(0.007)	
Lights MLD - average DN			0.018**			0.018**
			(0.008)			(0.008)
Lights standard deviation	0.002**	0.002***	0.002***	0.002**	0.002**	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unlit Grid Cells Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1636	1636	1636	1636	1636	1636
R-squared	0.024	0.032	0.033	0.024	0.032	0.034
Countries	51	51	51	51	51	51
Regional Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: (i) The outcome variable is gross cell product (i.e., total income) MLD from G-Econ data. (ii) Standard errors are clustered at the regional level. (iii) *DN* stands for digital numbers. (iv) The regressions are based on good quality gross cell G-Econ data for the years 1995, 2000, and 2005. (v) The unlit grid cells refer to the total grid cells with zero light intensity digital numbers (DN). (vi) This analysis includes 51 African Countries. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Figure 13: Regional Inequality Trends in Central Africa; 1992–2012



Source: Author’s calculations.

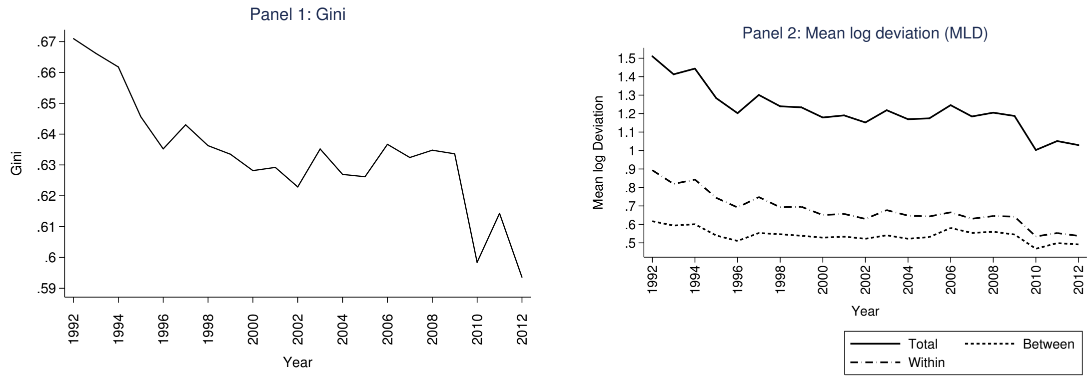
Figure 14: Regional Inequality Trends in Eastern Africa; 1992–2012

Source: Author’s calculations.

Figure 15: Regional Inequality Trends in Northern Africa; 1992–2012

Source: Author’s calculations.

Figure 16: Regional Inequality Trends in Southern Africa; 1992–2012

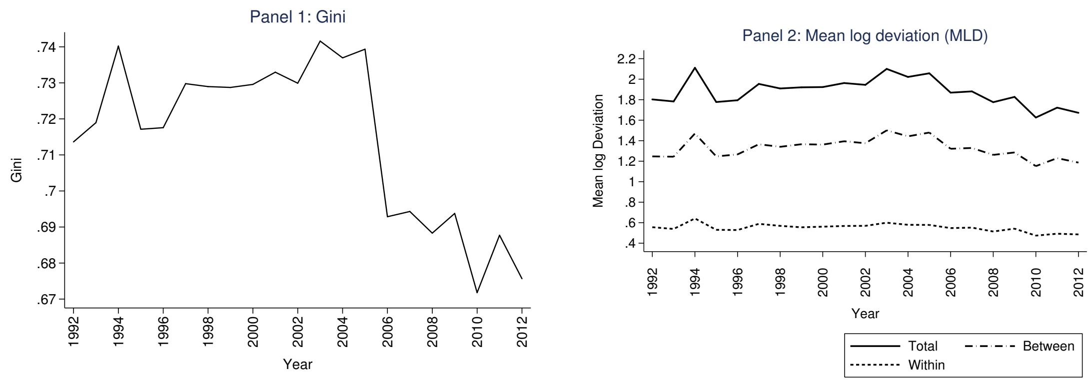


Source: Author's calculations.

Figure 17: Regional Inequality Trends in Western Africa; 1992–2012

Source: Author's calculations.

Figure 18: Regional Inequality Trends in Coastal Countries in Africa; 1992–2012



Source: Author's calculations.

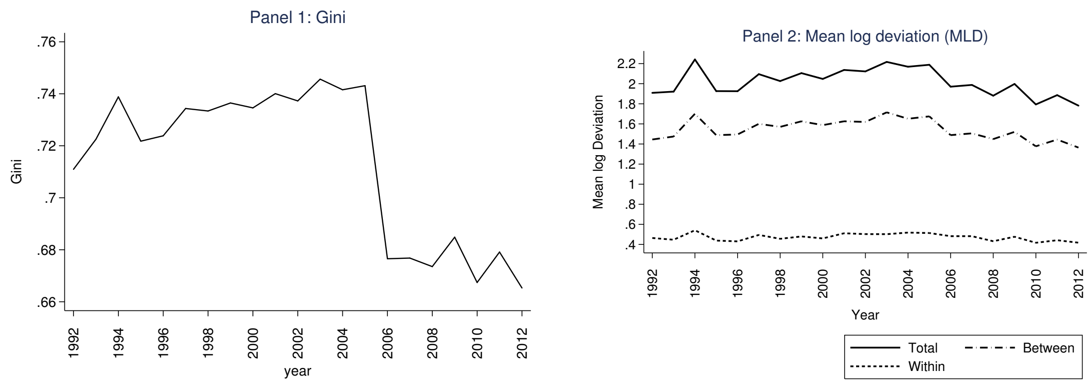
Figure 19: Regional Inequality Trends in Landlocked Countries in Africa; 1992–2012

Source: Author's calculations.

Figure 20: Regional Inequality Trends in Mineral Rich Countries in Africa; 1992–2012

Source: Author's calculations.

Figure 21: Regional Inequality Trends in Mineral Poor Countries in Africa; 1992–2012

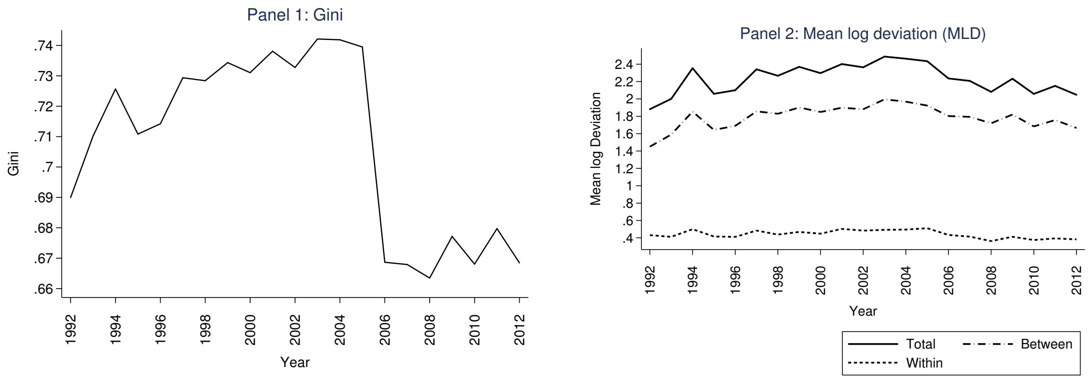


Source: Author's calculations.

Figure 22: Regional Inequality Trends in Countries Favourable to Agriculture in Africa; 1992–2012

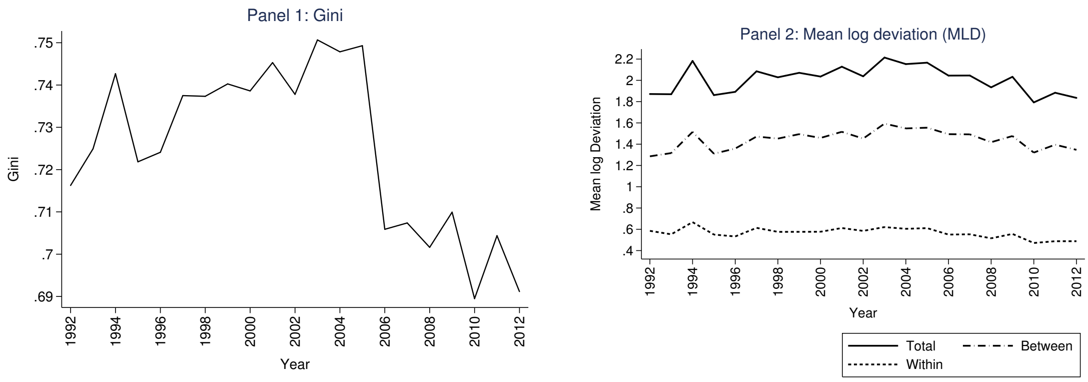
Source: Author's calculations.

Figure 23: Regional Inequality Trends in Countries Unfavourable to Agriculture in Africa; 1992–2012



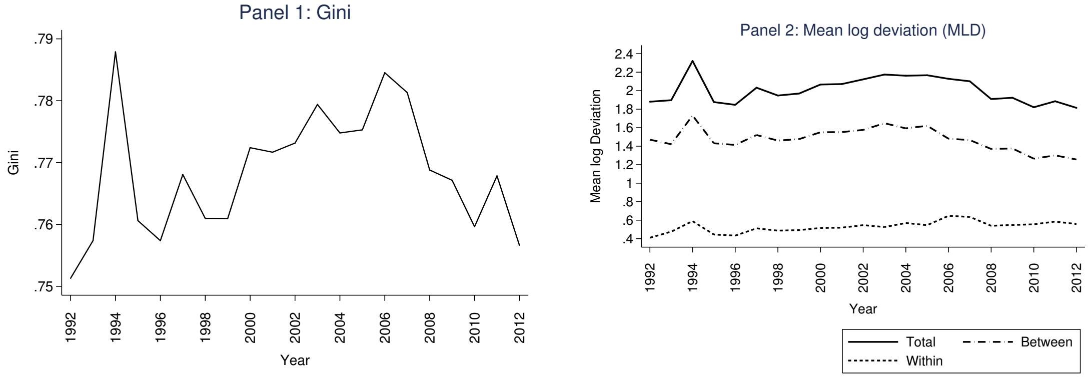
Source: Author's calculations.

Figure 24: Regional Inequality Trends in Former British Colony Countries in Africa; 1992–2012



Source: Author's calculations.

Figure 25: Regional Inequality Trends in Former French Colony Countries in Africa; 1992–2012



Source: Author's calculations.

Figure 26: Regional Inequality Trends in Former Portuguese Colony Countries in Africa; 1992–2012

Source: Author's calculations.

Figure 27: Regional Inequality Trends in Former Belgian Colony Countries in Africa; 1992–2012

Source: Author's calculations.