Top Lights

Bright cities and their contribution to economic development

Richard Bluhm[*]   Melanie Krause[†]

March 2022

Abstract

Tracking the development of cities in emerging economies is difficult with conventional data. This paper shows that nighttime lights can be used as a reliable proxy for economic activity at the city level, provided they are first corrected for top-coding. The commonly-used satellite images of nighttime light intensity fail to capture the true brightness of larger cities. We present a stylized model of urban luminosity and empirical evidence which both suggest that these 'top lights' can be characterized by a Pareto distribution or similarly heavy-tailed distributions. We then propose a correction procedure that recovers the full distribution of city lights. Our results show that the brightest cities account for nearly a third of global light output. Applying this approach to cities in Sub-Saharan Africa, we find that primate cities are outgrowing secondary cities. Contrary to the top-coded data, our data show that differences at the intensive margin drive the differential in relative growth rates across city types.

JEL Classification: O10, O18, R11, R12

Keywords: Development, urban growth, night lights, top-coding, inequality

[*]*Corresponding author.* Leibniz University Hannover, Institute of Macroeconomics, and University of California San Diego, Department of Political Science, e-mail: *bluhm@mak.uni-hannover.de*

[†]Leipzig University, Faculty of Economics and Management Sciences, Germany; E-mail: *melanie.krause@uni-leipzig.de*

1 Introduction

Cities are hubs of economic activity and productivity. About 4.2 billion people, or 55% of the world's population, currently live in urban areas, and developing countries are urbanizing at a rapid pace (United Nations, 2018). The African continent alone might add up to a billion people to its urban population by 2050. Many important questions in development economics and economic geography are intimately linked to understanding the processes driving the concentration of people and economic activity in cities. The lion's share of the extant literature is, however, oriented towards cities in advanced economies, for which ample data are available (Glaeser and Henderson, 2017). Much less is known about the rising cities of the 21st century, which not only lack comparable data but are undergoing such fast-paced urbanization that existing sources quickly become dated.

Satellite images of Earth are transforming how economists and other social scientists are tracking human activity and its consequences.[1] Nighttime images of light emissions are now an established proxy for local economic activity (Chen and Nordhaus, 2011, Henderson et al., 2012, Donaldson and Storeygard, 2016) and have been used in a variety of innovative applications (e.g. Michalopoulos and Papaioannou, 2013, Hodler and Raschky, 2014, Alesina et al., 2016, Pinkovskiy and Sala-i Martin, 2016, Lessmann and Seidel, 2017, Pinkovskiy, 2017, Henderson et al., 2018). They have several advantages over conventional survey-based data. Night lights are measured uniformly around the globe, allowing us to bypass discussions over adjustments for exchange rates and regional price levels. Moreover, the most widely-used data are available as an annual panel from 1992 to 2013 at a resolution of less than a square kilometer.

Our primary objective in this paper is to establish how these data can be used to reliably track economic activity within and across cities. A serious drawback of the standard night lights data is that they are top-coded in larger cities. The Operation Linescan System (OLS)—a part of the US Defense Meteorological Satellite Program (DMSP)—was designed to pick up dim light sources, but the satellites have a limited onboard storage capacity and are based on outdated 1970s technology. They record light intensities as integerized digital numbers from 0 DN (dark) to 63 DN (bright) and truncate all observations above this limit to save space. The upper end of this scale is, however, easily reached by the brightest lights emitted by a mid-sized city. As a result, the recorded signal "flatlines" when the satellites encounter bright city lights: the central business district and the outskirts of larger cities appear to be equally bright in the truncated data.[2]

---

[1]For reviews of the related literature see Donaldson and Storeygard (2016), who illustrate the advantages of remotely sensed data in general, and Michalopoulos and Papaioannou (2018), who focus on night lights data in particular.

[2]The data are known to suffer from a variety of other problems, such as bottom-coding (Jean et al.,

The scale of the truncation is sizable. In these so-called 'stable lights' data, the urban centers of large, busy cities, such as New York or London, appear to emit as little light as smaller American or British towns. Our estimates instead suggest that they are more than an order of magnitude brighter than recorded in the original data. While top-coding tends to affect developed countries more than their less-developed counterparts, we show that it is pervasive and distorts the ranking of cities within and across countries. Nearly all primary cities in Africa and mid-sized cities in Asia hit the top-coding threshold. Large agglomerations in developing countries, such as Johannesburg or New Delhi, are affected particularly strongly.

In this paper, we analyze the global distribution of city lights at the pixel level, develop a new procedure to recover the details of within city activity, and then study the evolution of cities in Sub-Saharan Africa. We make three distinct contributions to the literature:

First, we argue that it is natural to characterize the distribution of the world's brightest lights, which we dub 'top lights', by a Pareto distribution. We provide theoretical and empirical evidence supporting this claim. In terms of theory, we present a stylized model of light emissions from cities, combining standard assumptions on the evolution of city sizes (e.g., Zipf's law) with regularities in urban scaling. The model gives rise to a power law in light emissions above a certain threshold or other heavy-tailed distributions which are often approximated by a power law. Our empirical tests based on auxiliary satellite data also strongly favor a heavy-tailed Pareto distribution in top lights with an inequality parameter comparable to top incomes or wealth in the US (see, e.g., Atkinson et al., 2011).

Our second contribution is methodological. Building on the Pareto property of top lights, we develop a top-coding correction for the truncated data. To do so, we combine desirable features from the DMSP-OLS data with observations from the less frequently available 'radiance-calibrated' satellite data. Our correction procedure uses a geo-referenced ranking method at the pixel level. Based on this method, we present a new annual panel of nighttime lights over the entire period from 1992 to 2013.[3] Since 2012, nighttime lights are also observed by the technically superior Visible Infrared Imaging Radiometer Suite (VIIRS) which does not suffer from top-coding (see Gibson et al., 2021, for a recent comparison of both systems). Nevertheless, the DMSP data are still the benchmark series in economics, and our pixel-level panel for the 22-year period until 2013 should prove to be useful to researchers interested in studying the development of cities or other small geographies in the recent past. The DMSP-OLS satellite data have just been extended until 2018 and the system still records nighttime light intensities

---

2016), overglow or blooming (Abrahams et al., 2018), and geolocation errors (Tuttle et al., 2013).

[3]The corrected images can be downloaded from www.lightinequality.com. The data or versions of our correction approach are already being used in a number of applications (e.g. Jedwab et al., 2019, Pinkovskiy and Sala-i Martin, 2020, Gibson et al., 2021, Montalvo et al., 2021) as well as our own work (e.g. Düben and Krause, 2021, Castells-Quintana et al., 2021, Bluhm and McCord, 2022).

Figure 1 – Selected cities in 1999, stable lights and corrected lights

(a) New York City

(b) London

(c) New Delhi

(d) Johannesburg
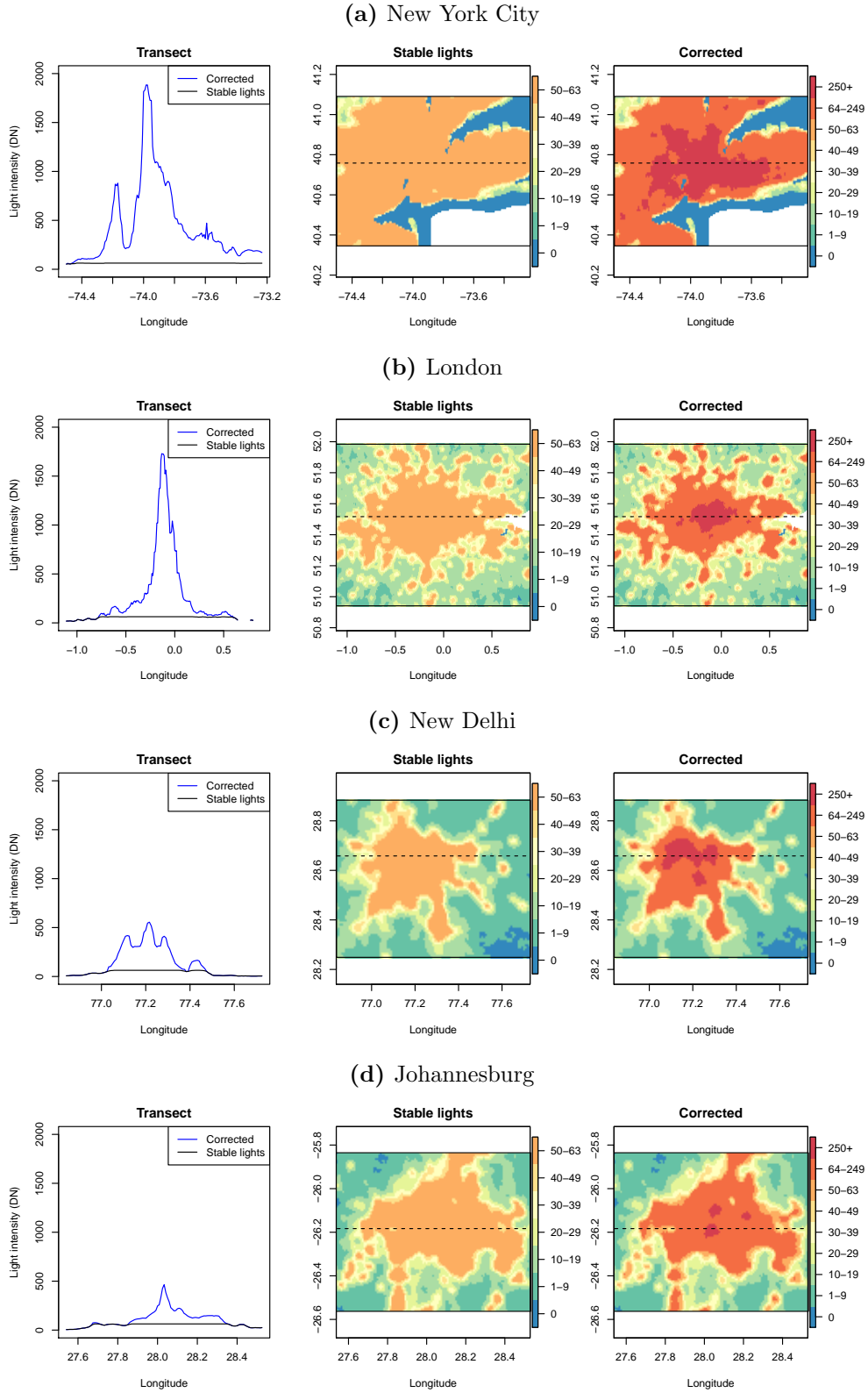
Notes: Comparison of the light intensities (DN) recorded by the stable lights and our corrected lights in four major cities in 1999. The left panel shows the light intensity along a longitudinal transect through the brightest pixel in each city. The middle panel shows a map based on the stable lights data from satellite F12 in 1999. The right panel shows the same map using the corrected data presented in this paper. Both data have been binned and the color scales were adjusted to be comparable. Dashed lines indicate the transect path.

during pre-dawn hours today. With some modifications, our correction can also be applied to this data or potential future extensions.[4]

We find that our correction makes a substantial difference in virtually all major cities. The top 4% of pixels—the average share of pixels which we correct globally—account for 32% of the total brightness observed on Earth. This is nearly double their original share, underlining the contribution of big cities to global economic activity. As a result, the spatial distribution of economic activity, as measured by nighttime lights, turns out to be much less equal.

Figure 1 illustrates the scale of the problem in the original data and compares these data with our solution. It shows that the stable lights data are unable to differentiate among the light intensities originating from different locations within New York, London, New Delhi, or Johannesburg. Across these four major cities, the average light intensity differs little and the sum of light in each city is largely driven by the urban extent. After our correction, we can clearly identify urban cores which are much brighter than the outskirts. The correction also matters for other subnational aggregates. The new data scale more linearly in economic density at the regional level in the United States or Germany than the original data.

Third, we use this new data to analyze the relative growth rates and internal structure of cities in Sub-Saharan Africa. The subcontinent is still the world's poorest region and urbanizing at a considerably lower level of development than other regions have in the past (Glaeser, 2014). Economic activity in Sub-Saharan Africa is heavily concentrated in primate and coastal cities due to a combination of factors, including high transport costs (Storeygard, 2016), outward-oriented colonial infrastructure (Jedwab and Moradi, 2016, Bonfatti and Poelhekke, 2017), urban bias (Lipton, 1977, Ades and Glaeser, 1995), the limited reach of national institutions (Michalopoulos and Papaioannou, 2014), climate change (Barrios et al., 2006), natural resource booms (Gollin et al., 2016), and the urban mortality transition (Jedwab et al., 2019). This structure is neither optimal nor static. On the one hand, excessive primacy in the city size distribution has been linked to the prevalence of slums and slow economic growth (Henderson, 2003, Castells-Quintana, 2017). On the other hand, recent studies suggest that Africa's secondary cities might be gaining ground vis-à-vis primate cities and playing an important role in poverty reduction (Henderson et al., 2012, Christiaensen and Todo, 2014, Christiaensen and Kanbur, 2017). Official numbers on economic activity at the city level remain elusive though so that it remains an open question if secondary cities in Africa are, in fact, rising.

Our application shows that Africa's largest cities are maintaining their dominant position. Primary cities in Sub-Saharan Africa were growing faster than secondary cities over the period from 1992 to 2013, suggesting they are continuously absorbing growing

---

[4]The radiance-calibrated data which serves as an auxiliary input to the correction only cover the period from 1996 to 2010, but a coarsened version of the VIIRS data could play the same role after 2012.

populations in informal settlements (in line with Jedwab et al., 2019). Working with the top-coding corrected data is crucial for understanding these developments and the results match the trends obtained from census-derived data on population at the city level. In fact, the uncorrected stable lights data give the impression that secondary cities grow as fast as primate cities during the period under investigation.

More broadly, this paper takes on the challenge of linking the pixel-level distribution of economic activity as observed from high-resolution satellites to economic theory and empirical regularities in urban economics. We argue that this approach leads to a better understanding of the features these data should have and at what scale they will be particularly useful. This perspective shows that the influence of top-coding is small when lights are aggregated to the country level but then rises steeply as the size of the unit of observation decreases. We also provide some evidence suggesting that a non-trivial part of the non-linearity in the relationship between light and GDP observed at higher economic densities arises from top-coding, rather than changes in the structural relationship. Moreover, as we demonstrate below, the impact of top-coding increases over time as countries and cities develop. Our application shows that even in Africa, which is still the poorest continent with the lowest electrification rates, top-coding makes a big difference. As the ultimate test of our argument, this suggests that top-coding will matter as least as much in other continents and regions. By providing a solution to this problem, we hope to further encourage researchers to use the historical nighttime lights series (and its extensions) in innovative ways.

The remainder of this paper is organized as follows. Section 2 provides some background on the nighttime lights data and the extent of top-coding around the world. In Section 3 we present theoretical and empirical evidence in favor of a Pareto distribution in top lights. Section 4 outlines our correction procedure. Section 5 contains the application of our top-coding corrected data to African cities. Section 6 concludes.

2 Top-coding around the world

The DMSP-OLS satellites have been orbiting Earth since the 1970s with the primary purpose of detecting clouds. As a byproduct, they measure night lights in the evening hours between 8:30 and 10:00 pm local time around the globe every day. Since 1992, the recorded data are pre-processed by the National Geophysical Data Center at the National Oceanic Administration Agency (NOAA) and averaged over cloud-free days.[5] The result is a set of images of annual 'stable light' intensities from 1992 to 2013 for every 30 by 30

---

[5]This removes observations of cloudy days and sources of lights which are not man-made, such as auroral lights or forest fires. The process also removes a lot of dim light sources. NOAA makes a series of unfiltered lights available, which we later use for the delineation of urban extents. Nighttime light observations at NOAA ended in 2019 and most of the group moved to the Payne Institute at the Colorado School of Mines which continues to produce annual, monthly, and daily products.

arc seconds pixel of the globe (about 0.86 square kilometers at the equator).[6]

We cannot use the truncated stable lights data to gauge the extent of top-coding. Fortunately, for seven years, the same satellites were sometimes operated with sensor settings that were less sensitive to light. NOAA generates a series of 'radiance-calibrated' lights by combining the stable lights data from normal flight operations with auxiliary data obtained from these low amplification sensors (Elvidge et al., 1999, Ziskin et al., 2010, Hsu et al., 2015). The resulting series is free of top-coding and has no theoretical upper bound.[7] The radiance-calibrated data can be used directly for cross-sectional analyses.[8] For comparisons over time, however, their sparse temporal coverage and fluctuations across years presents a significant drawback.

The radiance calibration process blends the frequency spectrum of several satellites into a single image but also introduces a substantial amount of noise. The resulting data are a mismeasured proxy of the true unobserved tail of the world's brightest lights. The variability is apparent in basic summary statistics. From 2004 to 2006, for example, global average luminosity falls by 14% and its standard deviation by 23% (Table A-1). Average luminosity in the tail fluctuates by 25% or more over the same period. Online Appendix A compares summary statistics of the stable lights and radiance-calibrated data over time. It also discusses technical reasons for the large fluctuations we observe in the latter. Due to these limitations, we rely on the shape of the distribution and the ranks of the radiance-calibrated pixels (instead of their actual values). In this manner, we combine desirable features from both the stable lights and the radiance-calibrated lights.

While the literature typically aggregates the pixel-level data to some study area of interest (e.g. grid cell, city, or region), we conduct our analysis at the native resolution of 30 arc seconds to avoid averaging over top-coded areas. On a global scale, this is a formidable task.[9] To ease the computational burden, we carry out large parts of the analysis with a spatial random sample of 10% of pixels within all countries that have a landmass larger than 500 km^2 (but later apply the correction to the full data). The sample contains more than two million pixels per year located in 194 countries or territories. Figure 2 provides a first visual insight into the lights at the top. It compares the distribution of the top 10% of the stable lights and the top 10% of radiance-calibrated lights in 2010. Top-coding is apparent in the (mirrored) J-shape of the former and the long tail of the latter indicates the possibility of a power law distribution.

---

[6]In our analysis, we always exclude areas close to the polar zones (65 degrees Southern and 75 degrees Northern latitude) known to be influenced by ephemeral lights and remove areas affected by gas flaring.

[7]Note that in spite of calibration issues, Hsu et al. (2015, p. 1865) point out that, within the same year, the "DNs below saturation of the Stable Lights product and DN_EQs of merged fixed-gain imagery can be directly compared to each other."

[8]Henderson et al. (2018), for example, use the 2010 version of this data in their study of the global spatial distribution of economic activity.

[9]Every image contains more than 700 million pixels, about a third of which are on land and half of which are lit.

Figure 2 – Histograms of nighttime lights in 2010, top 10% only

(a) Stable lights　　　　　　　　　　　**(b)** Radiance-calibrated lights
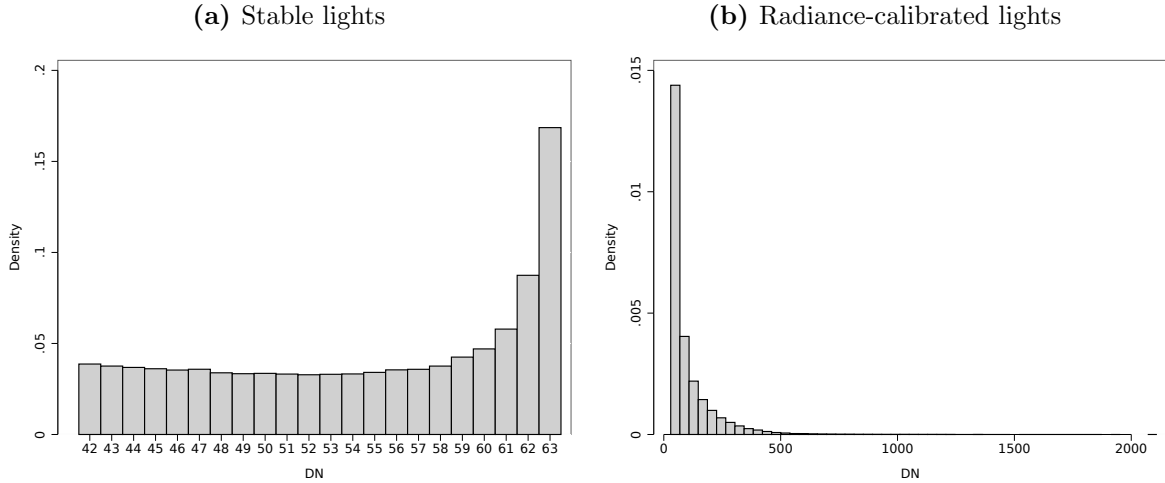
Notes: Histograms of the top 10% tail distributions of the stable lights (satellite F18), panel (a), and the radiance calibrated lights (satellite F16), panel (b). The input data are a 10% representative sample of all non-zero lights in the stable lights and radiance-calibrated data at the pixel level (see Elvidge et al., 2009, Hsu et al., 2015).

We still have to define where top-coding begins before we can assess its impact. Since the scale goes up to 63 DN, it seems natural to assume that this would be the appropriate threshold. There are, however, compelling technical and statistical arguments suggesting that the threshold should be much lower. Each value we observe in the annual data has been averaged several times. In addition to averaging the high-resolution data on-board of the satellites,[10] the daily observations are averaged again in the process of generating an annual image. Any pixel measured with a value of 63 on some, but not all, cloud-free days of a given year would have ended up with a yearly average below 63. Counting the share of pixels with the highest value, as is common practice, therefore only captures locations that are top-coded for the majority of cloud-free days of the year.[11]

Our analysis shows that many pixels with DNs of 62, 61, down to the mid-50s in the annual composites are subject to implicit top-coding and should be considerably brighter than the stable lights data suggest (see Online Appendix B). NOAA finds that top-coding's first–albeit faint–influence starts at much lower values, around 35 DN.[12]

---

[10]Abrahams et al. (2018) provide a detailed explanation of how the DMSP-OLS satellites process the data before transmitting them to Earth, including how this leads to geolocation errors, blurring, and top-coding. The satellite first truncates individual pixels at a much finer resolution and then aggregates those to a coarser resolution. Each night the origin shifts a bit, recreating a finer resolution. Finally, these data are aggregated and averaged again at NOAA when the annual composites are created.

[11]For example, a pixel would be top-coded at DN 63 in the annual average, if we observe a value of DN 63 in 51% of cloud-free days and a value of DN 62 in 49% of cloud-free days. However, this would no longer be true if the second value is lower than DN 62, in which case the effective threshold where top-coding plays a role begins to fall and more than 51% of days with DN 63 are required.

[12]Since "the OLS does onboard averaging to produce its global coverage data, saturation does not happen immediately when radiance reaches the maximum level. On the contrary, as the actual radiance grows, the observed DN value fails to follow the radiance growth linearly, causing a gradual transition

Yet, the true threshold is unknown. When settling on a particular value, we face a trade-off between overestimating the amount of top-coding if the chosen threshold is too low and underestimating if it is too high. Despite the suggestion that top-coding starts somewhere in the mid-30s, it does not dominate the empirical distribution until we consider substantially higher values. The left panel of Figure 2 shows that the density of recorded DNs falls until the mid-50s (as expected in the absence of top-coding) but at a decreasing rate and only then rises rapidly until the end of the scale at 63 DN. This pattern repeats itself in other years. We, therefore, suggest that a top-coding threshold of about 55 DN is a conservative choice that generates a decreasing histogram. We use this threshold throughout the paper but other options are possible.[13] In fact, our correction will work with any top-coding threshold that is a reasonable solution to this trade-off.

Table 1 – Summary statistics of global lights in 2010

	World	USA	Brazil	Israel	South Africa	China	Netherlands
	Panel a) Stable lights (from 0 to 63 DN)						
Mean	17.55	18.35	17.97	31.61	15.16	17.72	35.11
Standard deviation	15.35	16.90	15.45	21.49	15.29	15.50	16.09
Maximum	63.00	63.00	63.00	63.00	63.00	63.00	63.00
Spatial Gini	0.4258	0.4486	0.4165	0.3858	0.4604	0.4260	0.2637
Pixels	2,154,889	427,922	60,310	2,060	18,369	165,521	6,549
	Panel b) Radiance-calibrated lights (from 0 to ∞ DN)						
Mean	19.04	23.14	19.76	46.29	15.18	19.97	37.83
Standard deviation	44.35	53.42	41.76	82.70	29.99	46.71	44.08
Maximum	2109.67	1710.59	646.84	914.14	575.22	1862.04	435.63
Spatial Gini	0.6045	0.6613	0.5995	0.6505	0.5941	0.6093	0.4880
Pixels	2,154,889	427,922	60,310	2,060	18,369	165,521	6,549
	Panel c) Comparison of top-coded pixels						
Share = 63 DN	0.0176	0.0340	0.0228	0.1019	0.0206	0.0111	0.0299
Share $\geq$ 55 DN	0.0576	0.0847	0.0588	0.2447	0.0387	0.0597	0.1709
Radiance-cal. mean	142.05	153.77	143.36	135.70	113.27	143.96	107.92

Notes: The table reports summary statistics using the stable lights in Panel (a) and the radiance-calibrated lights in Panel (b). Panel (c) compares both sources at the pixel level. The top-coding shares are based on the stable lights data. The last line reports the radiance-calibrated mean of all pixels with saturated values $\geq$ 55 DN. All three panels are based on a 10% sample of all lit pixels, where each pixel is 30×30 arc seconds. The stable lights data are averaged across the whole year, while the radiance-calibrated data come from satellite F16, which recorded these data from January 11 to December 9, 2010.

Table 1 compares the stable and radiance-calibrated lights across the globe in 2010— the latest year where both data sources are available. The first column already shows that the difference in scale is striking. The brightest radiance-calibrated pixel in our sample

---

into a plateau of full saturation [...] this behavior can be observed starting at about DN = 35" (Hsu et al., 2015, p. 1872).

[13]The statistics in Table B-1 in Online Appendix B suggest that working with 56 or 57 rather than 55 would not have a big impact on our results, while the tail is much shorter and the histogram would have an unexpected shape when top-coding starts at 63 DN.

is more than 30 times brighter than the end of the scale of the stable lights. This is reflected in a standard deviation which is three times higher and a spatial Gini coefficient of inequality in lights which differs by 18 percentage points. In 2010, about six percent of all stable lights pixels are top-coded at 55 DN or higher, while their unsaturated counterparts are—on average—more than twice as bright.

Figure 3 – Share of top-coded pixels and country characteristics

(a) GDP per capita   **(b)** Log population density
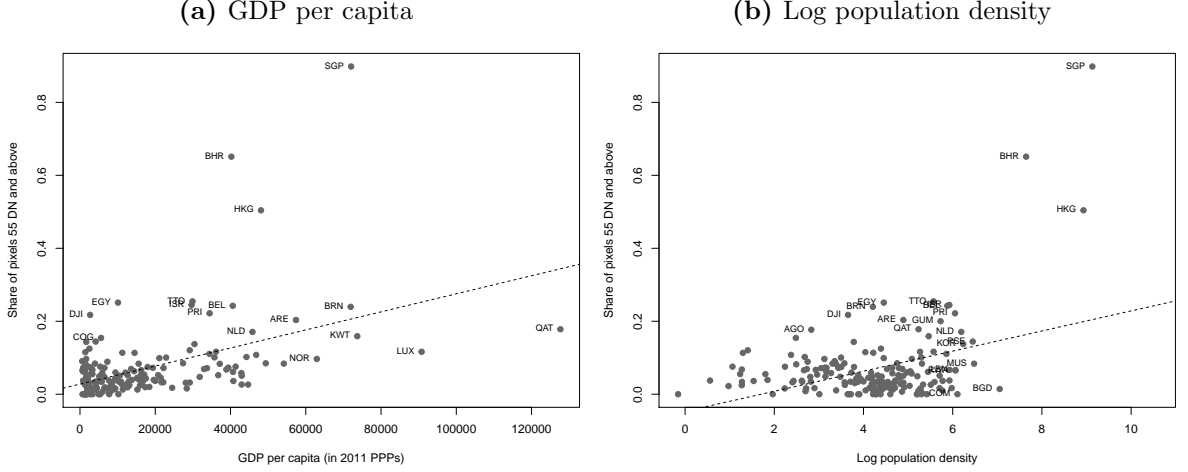
Notes: Illustration of the systematic bias introduced by top-coding. The data are a 10% representative sample of all non-zero lights in satellite F18 in 2010. GDP and population data are from the World Development Indicators.

Top-coding affects all countries but not uniformly. Countries that are *i)* highly developed, *ii)* small and *iii)* urbanized are more strongly affected by top-coding than others, but there is substantial heterogeneity. The remaining columns of Table 1 document this diversity in a selection of countries. Larger middle-income countries, like China, Brazil, or South Africa, have a top-coding share comparable to the world average. Mature economies of different sizes and population densities, such as the US, Israel, or the Netherlands, have greater top-coding shares from 8% up to 25%. In Israel and the Netherlands, a high average light intensity coupled with a high incidence of top-coding in the stable lights data generate such an artificially low spatial Gini coefficient that it rises by more than 20 percentage points in the radiance-calibrated data. Figure 3 illustrates these patterns across all countries in the sample and highlights the exceptions. For example, the overwhelming majority of pixels in high-income, high-density city-states, such as Singapore, Hong Kong, or Bahrain, are top-coded. Top-coding is also particularly pronounced in low-income countries with a low average population density but large primate cities, such as Egypt. Top-coding is thus a complex function of the spatial equilibrium, that is, the size, density, number, and location of cities in each country.

3 Investigating the distribution of top lights

Our correction approach rests on the claim that a Pareto distribution is a reasonable description of top lights. The Pareto distribution is an often-found empirical feature of data used in physics, biology, and many other sciences (see Newman, 2005), including the distribution of top income or wealth (e.g. Piketty, 2003, Atkinson et al., 2011) and the size distribution of cities as delineated via night lights (Small et al., 2011). So far it has not been established for the upper tail of the distribution of nighttime light intensities.

The main issue we face is that the true tail of the light distribution is unobserved. As a result, we cannot directly prove that a Pareto distribution is "correct." We support our claim that the Pareto distribution is a reasonable working assumption in two different ways. First, we present a tractable model of light intensities within and across cities, showing that the density of top lights at the pixel level under given assumptions is either Pareto or leads to heavy-tailed distributions which are often approximated by a Pareto distribution. Particular assumptions of the model, such as monocentricity, differ from reality and are unnecessary for the correction procedure we propose later on. In the second part of this section, we make the empirical case in favor of a power law in top lights by examining the radiance-calibrated data, which we consider an imperfect proxy for the unobserved tail, and subjecting them to a battery of tests.

3.1 A stylized model of city lights

Our starting point is the size distribution of cities in terms of their population. It is well known that combining Gibrat's (1931) law of homogeneous growth of cities with a lower bound of city sizes leads to a Pareto or Zipf distribution of large cities (Gabaix, 1999, Eeckhout, 2004). Zipf's law predicts that city ranks are inversely proportional to their size; for instance, the biggest city in the U.S. (New York) has twice the population of the second-ranked city (Los Angeles) and three times the population of the third-ranked city (Chicago).

Numerous studies have provided empirical evidence for this regularity based on U.S. cities or metropolitan areas (e.g. Gabaix and Ioannides, 2004). Despite heterogeneity in the rank-size parameter in other countries (Rosen and Resnick, 1980, Soo, 2005), it is generally considered a good approximation of the size distribution of large cities (Luckstead and Devadoss, 2014). Evidence in its favor becomes even stronger when cities are defined by agglomerations as opposed to administrative boundaries (e.g. see Rozenfeld et al., 2011).

Assumption 1. *The size distribution of big cities in terms of their population x is Zipf, i.e. a Pareto distribution with shape parameter $\eta = 1$ above some threshold x_c.*

The CDF of the number of cities with population x is

$$F(x) = 1 - \left(\frac{x_c}{x}\right)^{\eta} = \int_{x_c}^{x} \eta \frac{x_c^{\eta}}{x^{\eta+1}} \mathrm{d}x = \int_{x_c}^{x} \frac{x_c}{x^2} \mathrm{d}x. \tag{1}$$

Cities have a certain area extent in which they house their population. The urban allometry literature (Stewart, 1947, Jones, 1975), which deals with the scaling of human-made structures within cities, links the two quantities in the following way:

Assumption 2. *The population x and the area s of a city are related by a power law, i.e. $x \sim s^{\phi}$.*

It is typically assumed that $1 < \phi < 2$, so that larger cities not only spread out on the plane by converting the surrounding agricultural land but also grow in terms of height (Batty and Longley, 1994). Bettencourt (2013), for example, motivates scaling laws based on a network theory of human interactions using the parameter value $\phi = 1.5$.

To derive the distribution of individual pixels y within cities, we model the shape of cities. The standard assumption in the workhorse model of urban economics is monocentricity (Mills, 1967, Desmet and Rossi-Hansberg, 2013).

Assumption 3. *Cities are monocentric and of circular shape. They consist of rings of unit width from the center to the outskirts.*

Depending on its area s and therefore its population x, each city has $r = \pi^{-1/2} x^{1/(2\phi)}$ rings.[14] Integration by substitution yields the CDF of the number of rings per city[15]

$$F(r) = 2\phi x_c \pi^{-\phi} \int_{\tilde{r}}^{r} r^{-2\phi-1} \mathrm{d}r = \begin{cases} 0 \text{ for } r < \tilde{r} \\ 1 - x_c \pi^{-\phi} r^{-2\phi} \text{ for } r >= \tilde{r} \end{cases} \tag{2}$$

with $\tilde{r} = \pi^{-1/2} x_c^{1/(2\phi)}$. Its density

$$f(r) = x_c \pi^{-\phi} r^{1-2\phi} \text{ for } r >= \tilde{r} \tag{3}$$

follows a power law with a shape parameter $2\phi - 1 > \eta = 1$ for $\phi > 1$. Note that the shape parameter is higher than in the distribution of city sizes in terms of their population. This implies that the distribution of rings has fewer extreme values.

Each city with r rings consists of πr^2 rectangular pixels of unit size. The pixels are located at distance d from their respective city center. There are two opposing effects governing how the global number of pixels depends on this distance: *i)* within a given city, the number of pixels increases linearly with d because rings farther from the center

---

[14]For simplicity of exposition, we assume $x = s^{\phi}$, suppressing the proportionality constant.
[15]Online Appendix C derives this result.

contain more pixels, and *ii)* the larger the distance d from the city center, the fewer cities of that size are left, namely only the cities with $r \geq d$ rings.

Dividing the absolute amount of pixels at each distance d by the total number of pixels yields their global density

$$f(d) = \begin{cases} 2\pi\frac{\phi-1}{\phi}x_c^{-1/\phi}d & \text{for } d < \tilde{d} \\ 2\pi^{1-\phi}\frac{\phi-1}{\phi}x_c^{1-1/\phi}d^{1-2\phi} & \text{for } d \geq \tilde{d} \end{cases} \tag{4}$$

with $\tilde{d} = \pi^{-1/2}x_c^{1/(2\phi)}$.[16] This density is illustrated in the left panel of Figure 4.

To derive the global density of luminosity $f(l)$, we require a last assumption on how light intensity within a city varies with the distance to the center d. As the within-city light distribution has not yet been modeled explicitly, we resort to standard models of population density p.

A popular choice in the literature is the negative exponential function, that is, $p(d) = P_0 \exp(-\gamma d)$ where $p(d)$ is the population density at distance d from the city center, $P_0 \geq p$ is the density at the center, and $\gamma > 0$ is a decay parameter so that the city periphery is more sparsely populated or, in our case, lit. The negative exponential can be motivated on the basis of the standard Alonso-Muth-Mills model (Brueckner, 1982). Bertaud and Malpezzi (2014) test it empirically on a large sample of cities, finding an average $\gamma = 0.1563$ for cities where the negative exponential fits well (R^2 above 0.9).

An alternative to the exponential distribution for population is the inverse power function, which was originally proposed for gravity models of traffic flow (Smeed, 1961, Coleman, 1964, Batty and Longley, 1994). It is defined as $p(d) = P_0 d^{-a}$ with $d > 0$ and shape parameter $a > 0$. Both functions are qualitatively similar for intermediate distances, with the inverse power function slightly below the negative exponential. They differ from each other in the city center, where the inverse power function has higher and more sharply declining values, as well as in the outskirts, where it predicts a higher density (see the right panel of Figure 4).

Turning towards lights, we note that light intensity in a pixel is the analog of the product of population and income per capita, and differences in light intensity within countries are mostly driven by variations in population density (Henderson et al., 2018). We assume that light is proportional to population within cities and, hence, can be fully characterized by the population density gradient.

Assumption 4. *Within cities, lights l are proportional to population p.*

We have no priors on whether the within-city light distribution can best be described by the inverse power function or the negative exponential.[17] We discuss the implications of both possibilities, starting with the case of the inverse power function.

---

[16]Online Appendix C derives this result.

[17]There are arguments suggesting that lights could follow either one of these distributions more closely.

Figure 4 – Illustration of distributions

(a) Pixel density

(b) Light gradient
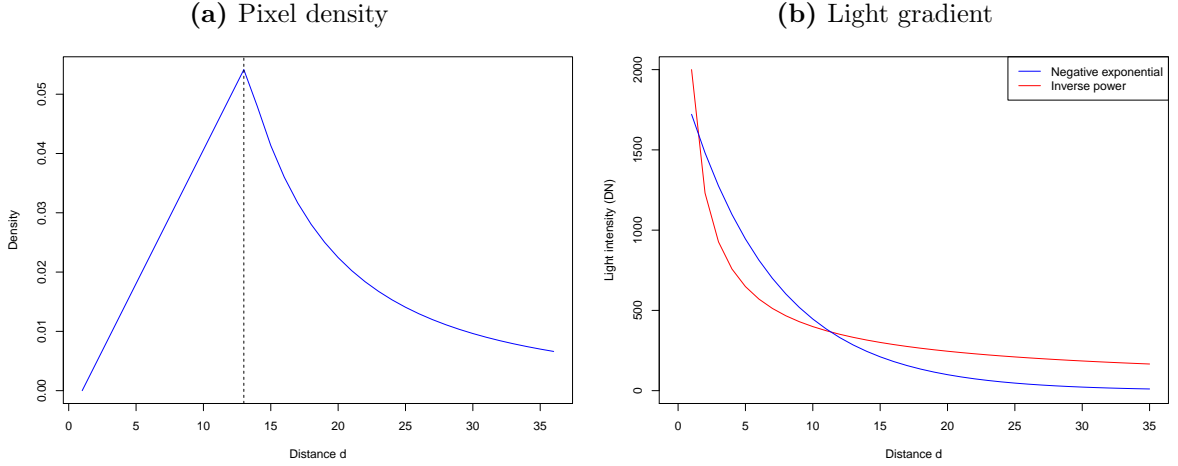
Notes: The left panel shows the pixel density from eq. (4) with $x_c = 10000$ and $\phi = 1.5$. The right panel shows the negative exponential distribution with $\gamma = 0.15$ as well as the inverse power distribution with $a = 0.7$; both start at $P_0 = 2000$.

Assumption 5a. *Within cities, light density $l(d)$ follows an inverse power function $l(d) = L_0 d^{-a}$ for $d > 0$ with $L_0 > l$ as the maximum luminosity at the center and $a > 0$ as the decay parameter.*

Applying the variable transformation from the inverse power function to the pixel density in eq. (4) yields

$$
f(l) = \begin{cases} 2\pi^{1-\phi}\dfrac{\phi-1}{\phi}x_c^{1-1/\phi}\Big(L_0/l\Big)^{(1-2\phi)/a} & \text{for } 0 < l \leq \tilde{l} \\[2mm] \underbrace{2\pi\dfrac{\phi-1}{\phi}x_c^{-1/\phi}}_{c}\Big(L_0/l\Big)^{1/a} & \text{for } \tilde{l} < l \leq L_0, \end{cases}
\tag{5}
$$

where $\tilde{l} = L_0 \pi^{a/2} x_c^{-a/(2\phi)}$.

Restricting our attention to the upper part of the light distribution from $\tilde{l}$ onwards, we can establish our key result:

Result 1a. *Based on assumptions 1–4a, top lights above threshold $\tilde{l}$ follow the Pareto distribution $f(l) = c\Big(L_0/l\Big)^{1/a}$ with shape parameter $1/a$.*

Next, we derive the distribution of top lights when lights within a city follow a negative exponential function instead of an inverse power function. Replace assumption 5a by

---

For example, due to its greater concentration at the center, the inverse power function is known to be particularly suitable for business floor-space models (as recommended by Zielinski, 1980), which is likely to be reflected in the distribution of lights as well. The tail of the inverse power function also typically fits the urban fringe well (Parr, 1985). On the other hand, the negative exponential distribution is both very flexible and more widely-used in empirically modelling population densities (Bertaud and Malpezzi, 2014). If there is proportionality between population and light, then lights would follow this distribution closely.

14

Assumption 5b. *Within cities, light density $l(d)$ follows an exponential function $l(d) = L_0 \exp(-\gamma d)$, where $L_0 \geq l$ is the density at the center and $\gamma > 0$ is a decay parameter.*

This set-up does not directly generate a Pareto distribution in lights. The distribution now depends on L_0, the maximum luminosity of the center of each city. For the light density $f(l)$ conditional on L_0, we use the variable transformation of eq. (4) together with $d = \frac{1}{\gamma} \ln(L_0/l)$ to obtain

$$f(l \mid L_0) = \begin{cases} 2\pi^{1-\phi} \frac{\phi-1}{\phi} x_c^{1-1/\phi} \left[\frac{1}{\gamma} \ln(L_0/l)\right]^{1-2\phi} \text{ for } l \leq \tilde{l} \\ \underbrace{2\pi \frac{\phi-1}{\phi} x_c^{-1/\phi} \frac{1}{\gamma}}_{c} \ln(L_0/l) \text{ for } \tilde{l} < l < L_0, \end{cases} \tag{6}$$

where $\tilde{l} = L_0 \exp(-\gamma \tilde{d})$.

This conditional density increases for dim light intensities (at the fringes of the largest cities) and decreases from $\tilde{l}$ onwards. The turning point and maximum, $\tilde{l}$, corresponds to the minimum size, $\tilde{d}$, of each city in terms of distance from the center. At higher light intensities, there are fewer and fewer pixels as these are the ones located in ever-smaller rings closer to the center.

To derive the marginal density $f(l)$, we have to make additional assumptions about the distribution of maximum light intensities, L_0, across cities. We distinguish among three cases, ranging from the most conservative one (constant L_0) over an intermediate case (L_0 follows a uniform distribution) to another power law (L_0 follows a Pareto distribution across cities). To benchmark the plausibility of these three cases, Figure 5 provides a scatter plot of the maximum light intensities of 988 world cities (based on the 2010 radiance calibrated satellite) over city sizes and plots the distribution of the maximum light intensities.

Case 1: L_0 is a constant, so that all cities large and small are equally bright in the center. This case is trivial but unrealistic given Figure 5. It is immediately obvious that no power law emerges in the distribution of top lights.[18]

Case 2: The maximum luminosity L_0 varies across cities. We opt for the most conservative specification and let L_0 follow a uniform distribution between values L_{min} and L_{max} with density $f(L_0) = (L_{max} - L_{min})^{-1}$. Cities big and small differ in their maximum luminosity, but all maxima are equally likely across cities. If we ignore the positive slope, Figure 5 suggests that there is some evidence for this assumption given the large range of maximum light intensities at a given city size. We can now write the

---

[18]The marginal density of lights $f(l)$ equals the conditional density $f(l \mid L_0)$. Analyzing the top end of the distribution $c \ln(L_0/l)$ for $\tilde{l} < l < L_0$, we observe a near linear decrease with a slope of $c \frac{d}{dl} \ln(L_0/l) \propto -\frac{1}{l}$ for l close to L_0. There is no power law because the assumption of a constant L_0 across all cities generates too many pixels with the highest luminosities.

joint density of l and L_0 as

$$f(l, L_0) = f(L_0)f(l \mid L_0) = \frac{c}{L_{max} - L_{min}} \ln(L_0/l) \text{ for } L_0 > l. \tag{7}$$

The marginal density, $f(l)$, is found by integrating out L_0

$$f(l) = \frac{c}{L_{max} - L_{min}} \int_l^{L_{max}} \ln(L_0/l)dL_0 = \frac{c}{L_{max} - L_{min}} \left[L_0(\ln L_0 - \ln l - 1)\right]_l^{L_{max}}$$
$$= \frac{c}{L_{max} - L_{min}} \left[l + L_{max}(\ln(L_{max}/l) - 1)\right], \tag{8}$$

which holds for $\tilde{l}_{max} < l < L_{max}$.

Case 3: Some city centers are much brighter than others. We now assume that L_0 follows a Pareto distribution across cities. Most of our empirical evidence points in this direction. The histogram in Figure 5 suggests the possibility of a power law in the tail, which is supported by additional empirical tests.[19] Note that we are only interested in the upper part of the density in eq. (6).[20] If L_0 follows the Pareto density with $\eta = 1$ so that $f(L_0) = L_{min}/L_0^2$, with L_{min} as the minimum center luminosity, the joint density of

Figure 5 – Distribution of maximum light intensity, L_0, across 988 world cities

(a) Scatter plot

(b) Histogram
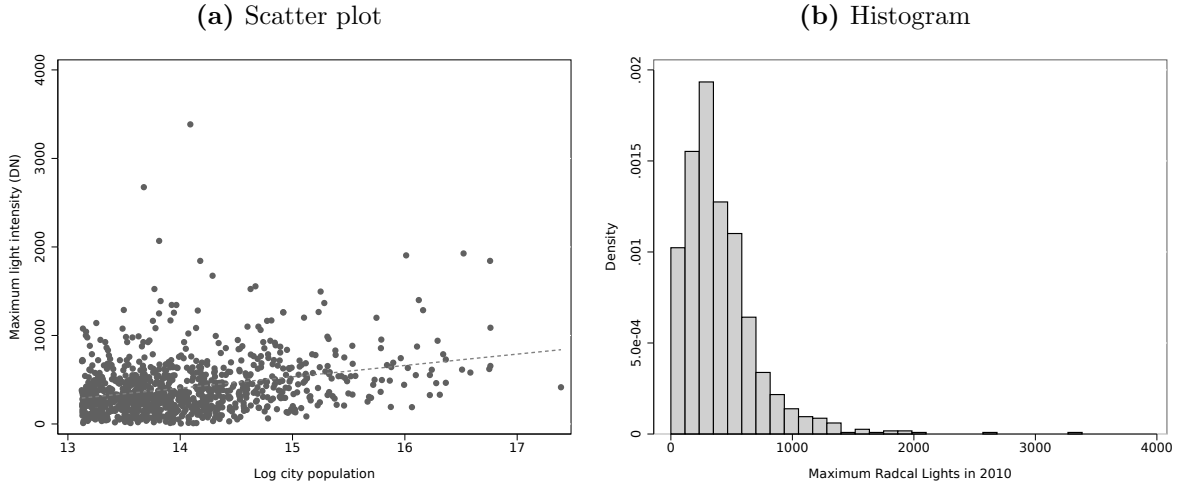
Notes: The plot shows the maximum light intensity recorded by the 2010 radiance calibrated satellites within a 25 km radius of the city center of 988 world cities with more than 500,000 inhabitants.

---

[19]In Appendix Online Appendix C, we conduct various tests for whether the distribution of the maxima is Pareto. No matter which threshold we choose, a discriminant moment ratio plot (Cirillo, 2013) suggests the data are Pareto distributed. Log-rank regressions using either the OLS or Hill estimator imply a Pareto parameter of around 3.

[20]The threshold $\tilde{l}$ depends on the random variable L_0. Hence, we restrict our analysis to the area of the density starting at the highest possible threshold value $\tilde{l}_{max}$, corresponding to $L_0 = L_{max}$.

l and L_0 is

$$f(l, L_0) = f(L_0)f(l \mid L_0) = c\frac{L_{min}}{L_0^2}\ln(L_0/l) \text{ for } \tilde{l}_{max} < l < L_{max} \text{ and } L_0 > l. \tag{9}$$

The marginal density, $f(l)$, is again found by integrating out L_0

$$
\begin{aligned}
f(l) &= cL_{min}\int_l^{L_{max}}\frac{1}{L_0^2}\ln(L_0/l)dL_0 = cL_{min}\left[\frac{\ln l - \ln L_0 - 1}{L_0}\right]_l^{L_{max}} \\
&= cL_{min}\left[\frac{1}{l} - \frac{1}{L_{max}}(\ln(L_{max}/l) + 1)\right],
\end{aligned}
\tag{10}
$$

which holds for $\tilde{l}_{max} < l < L_{max}$.

Result 1b. *Based on Assumptions 1–3 and 4b, as well as a sufficient variation in maximum light intensity across cities (case 2 and 3), top lights above a threshold $\tilde{l}$ follow a heavy-tailed distribution.*

The heavy-tailed distributions in eq. (8) and eq. (10) differ mathematically from the simple Pareto. However, it is common practice in the applied literature to approximate such heavy-tailed distributions by a Pareto distribution (Cowell and Flachaire, 2007, Nordhaus, 2012). Figure 6 illustrates how the Pareto distribution fits these distributions for selected parameters ($\alpha = 1.5$ for case 3 and $\alpha = 1.2$ for case 2).

Figure 6 – Approximating the theoretical densities with Pareto distributions

(a) Case 2 and Pareto with $\alpha = 1.2$ **(b)** Case 3 and Pareto with $\alpha = 1.5$
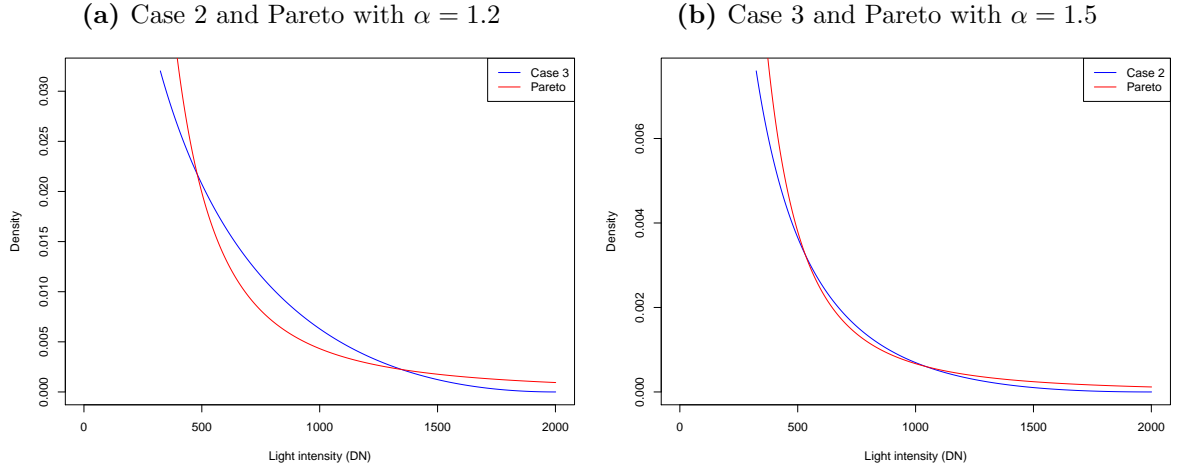
Notes: The left panel shows the pixel density from eq. (8) with $x_c = 10000$, $\phi = 1.5$ and $\gamma = 1.5$. The Pareto distribution with $\alpha = 1.2$ is scaled to fit. The right panel shows the pixel density from eq. (10) with $x_c = 10000$, $\phi = 1.5$ and $\gamma = 1.5$. The Pareto distribution with $\alpha = 1.5$ is scaled to fit.

No matter if the within-city light distribution follows an inverse power function or a negative exponential distribution with cross-city maximum light variation, we either obtain a Pareto distribution in top lights or a heavy-tailed distribution that can be

approximated by a Pareto distribution. Using a grid-search, we find that for $a = 0.7$, the inverse power function comes closest to the negative exponential function of the empirical parameter $\gamma = 0.15$.[21] A back-of-the-envelope calculation thus implies a Pareto shape parameter around one and a half. Note that we will econometrically estimate this parameter using auxiliary data in the next subsection.

The main theoretical insight is that the model can generate a Pareto distribution in top lights based on limited assumptions. If the within-city gradient of lights resembles that of floor space, then the theoretical distribution of top lights is Pareto. If not, then we still obtain heavy-tailed distributions. The correction we propose later relies on the Pareto property and does not require any other simplifying assumptions, such as monocentricity or a particular light density gradient. One possible application of our corrected data is to study how sub-centers emerge within larger cities.

3.2 Empirical tests

Empirical tests for a Pareto tail were popularized in economics by the literature on top incomes (e.g. Piketty, 2003, Atkinson et al., 2011) and city sizes (e.g. Rosen and Resnick, 1980, Gabaix, 1999, Gabaix and Ioannides, 2004, Rozenfeld et al., 2011). These tests usually exploit particular properties of the Pareto distribution.

Recall that data y which is Pareto distributed above a certain threshold y_c has the probability density function, $f(y) = \alpha y_c^\alpha y^{-\alpha-1}$, where α is the relevant shape parameter and only takes on positive values. The survival function, $1 - F(y) = (y_c/y)^\alpha$, gives the probability that the random variable Y is larger than the given value y. This maps directly to our model from the previous section, only that we now denote the light intensity at the pixel level by y and the shape parameter by $\alpha = 1/a$ to simplify the exposition.

For these tests, it is important to keep in mind that the true tail distribution in the stable lights data is, by definition, unobserved. Yet, with some caveats, the radiance-calibrated data and the newer VIIRS data are informative about the shape of the distribution and the relative brightness of individual pixels.

Visual inspection: Following Cirillo (2013), we visually check whether our data are Pareto distributed, before estimating shape parameters that are only meaningful if this condition is fulfilled. We use the seven satellite-years of radiance-calibrated data to make inferences about the shape of the missing tail in the stable lights data. Figure 7 shows a discriminant moment ratio plot with the coordinate pair of the coefficient of variation (i.e., standard deviation divided by the mean) on the x-axis and skewness on the y-axis (Cirillo, 2013). Each parametric distribution has its particular curve of feasible coordinates so that

---

[21]The curve-fitting was conducted on the domain for d from 1 to 50, minimizing the squared error between the negative exponential function with $\gamma = 0.15$ and the inverse power function with $a \in [0.05, 2]$.

the relevant part of the plane can be divided into a Pareto area (comprising Pareto type I and type II distributions), a log-normal area, and a gray area possibly belonging to both. This type of plot provides a more reliable indication of the Paretian nature of the data than more traditional graphical devices (such as Zipf plots shown in Online Appendix E).

Figure 7 – Visual evidence

(a) DMR plot, top 4%

(b) DMR plot, top 1%

(c) Histogram and fit, top 4% in 2010

(d) Histogram and fit, top 1% in 2010
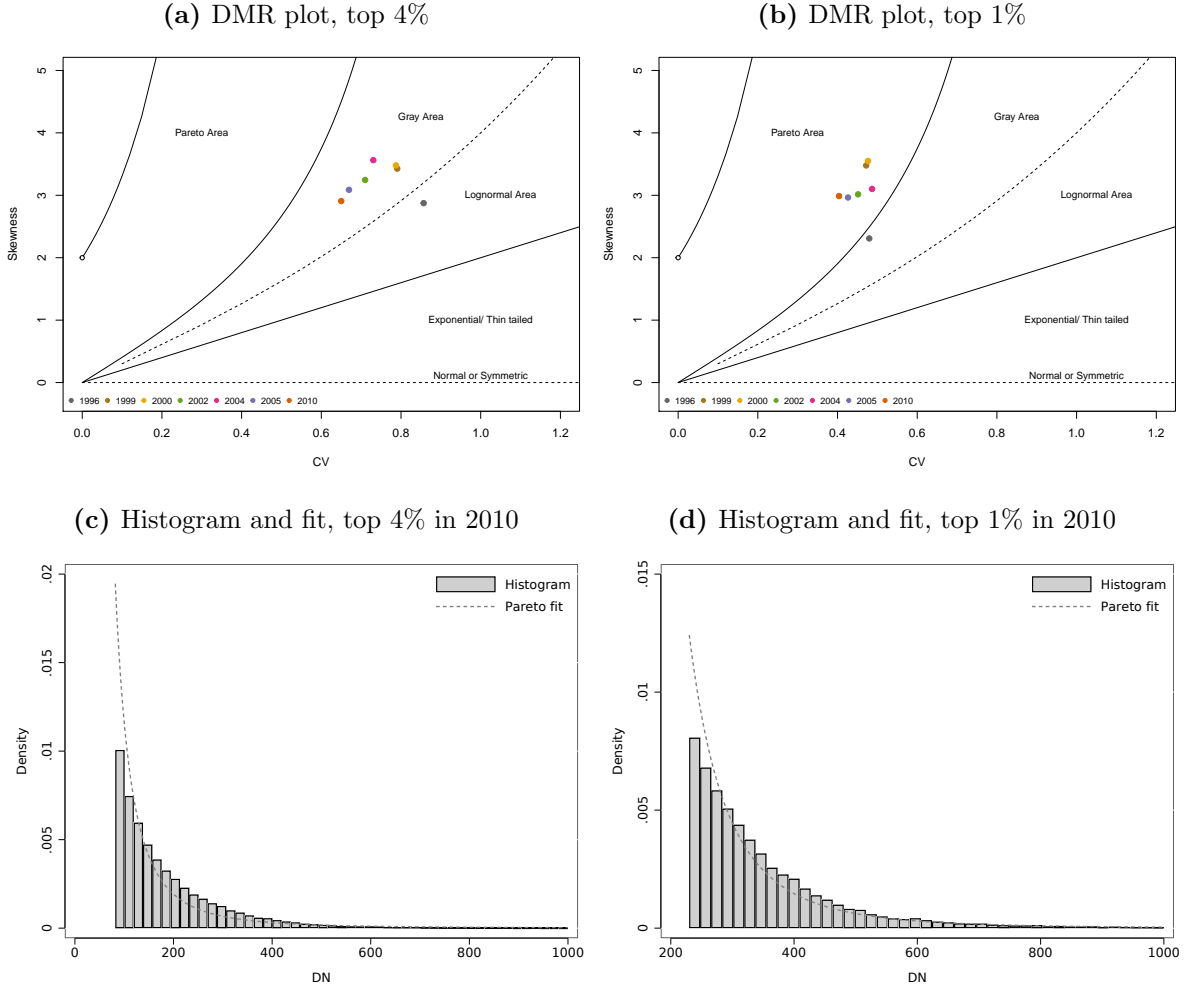
Notes: The figures in panels (a) and (b) show discriminant moment ratio plots (Cirillo, 2013), while the figures in panels (c) and (d) show histograms of the tail distribution of the radiance-calibrated lights in 2010 and a fitted Pareto distribution (with the shape parameter corresponding to the percentile, this is 1.5963 for the top 4% and 2.8641 for the top 1%, see Table 2). The histograms zoom in on the values below 1000 DN for the ease of exposition. The input data are a 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level, where each pixel is 30×30 arc seconds.

The visual evidence in favor of a Pareto (type I or II) distribution is strongest for higher thresholds. For the top 4% of pixels in panel (a), all but one satellite are located in the area of indeterminacy.[22] This ambiguity mirrors similar findings in the city-size

---

[22]The only exception is satellite F16 in 1996 which is based on considerably fewer cloud-free overpasses than later years. It is the earliest and dimmest of the radiance-calibrated products. Its highest values are far below those in subsequent years and the data contain many ties. Online Appendix A contains

literature, where both the Pareto and the log-normal with a large standard deviation generate tails that are virtually indistinguishable (Eeckhout, 2004, 2009).[23] For smaller percentages, such as the top 1% in panel (b), the evidence in favor of a Pareto-type distribution becomes much stronger. All but one satellite are in the Pareto area, far from the log-normal area and those of other thin-tailed distributions. These insights are corroborated by the histograms of the top 4% and top 1% in Panels (c) and (d) in Figure 7, which are overlaid with a Pareto distribution of the fitted parameters.

Online Appendix E examines the issue of log-normal versus Pareto in more detail. We conduct several goodness-of-fit tests, where we compare the empirical distribution of the radiance-calibrated data to the best-fitting theoretical distribution. The findings are unambiguous. The Pareto distribution mimics the distribution of top lights much better than the log-normal (with R^2s close to unity). Hence, we conclude that our data are best characterized by a heavy-tailed Pareto-like distribution. For robustness, we consider the truncation purely as a statistical issue of tail probabilities and use extreme value theory to find their distribution. Reassuringly, this approach also points towards a Paretian distribution of top lights (see Online Appendix D).

Of course, the heavy-tailed distributions found here can be (perhaps even more closely) approximated by a range of flexible functions (including Pareto distributions with multiple parameters, general power laws, or generalized Pareto curves). Non-parametric approaches could also be used if we abstract from the model and only consider the two data sources that we use in the correction. For parsimony, we restrict ourselves to the simple (type I) Pareto throughout this paper and use this parametric representation of the unobserved tail to minimize the noise inherent in the radiance-calibrated data.

Log-rank regressions: Log-rank regressions are a popular approach to estimating the Pareto shape parameter and firmly rooted in urban economics. They are based on the following approximation. For Pareto-distributed observations y_i, $i = 1, ... N$, with the survival function given above, we have $\text{rank}(y_i) \approx N y_c^\alpha y_i^{-\alpha}$, or, in logarithms $\log \text{rank}(y_i) - \log N \approx \alpha \log y_c - \alpha \log y_i$. Gabaix and Ibragimov (2011) show that directly estimating this relationship systematically underestimates the true coefficient and standard error. However, once the ranks are shifted by minus one-half and the standard errors are adjusted, rank regressions consistently estimate the parameters of

---

the relevant summary statistics for the radiance-calibrated data.

[23]It is well-established in the literature on city sizes and top incomes that the Pareto distribution is often only a good representation for the very top of data, with a moderately decreasing fit as the threshold decreases. Our data suggest that a Pareto distribution of top lights remains a good approximation down to the top 10% of the data. This is more than sufficient for our purposes, since we only consider 3% to 5% of all pixels to be top-coded in any given year. Indeterminacy vis-à-vis the log-normal is common in this literature as well, since "the tail of a log-normal is indistinguishable from the Pareto under certain circumstances, [so that] the researcher who is interested in the tail properties of a size distribution can choose which one to use" (Eeckhout, 2009).

interest and turn out to be relatively robust to deviations from power laws.

Table 2 – OLS rank regressions

Year	1996	1999	2000	2003	2004	2006	2010	Average
				Panel a) Top 5%				
Pareto $\hat{\alpha}$	1.3122	1.2993	1.2471	1.4482	1.4484	1.4790	1.4453	1.3828
	(0.0060)	(0.0054)	(0.0054)	(0.0065)	(0.0063)	(0.0066)	(0.0062)	[0.0932]
Observations	96,685	116,858	106,914	100,095	106,899	99,487	107,745	–
				Panel b) Top 4%				
Pareto $\hat{\alpha}$	1.3514	1.4000	1.3594	1.5890	1.5837	1.6250	1.5963	1.5007
	(0.0069)	(0.0065)	(0.0066)	(0.0079)	(0.0077)	(0.0081)	(0.0077)	[0.1236]
Observations	77,348	93,484	85,482	80,075	85,489	79,590	86,196	–
				Panel c) Top 3%				
Pareto $\hat{\alpha}$	1.4270	1.5665	1.5444	1.7667	1.7876	1.8423	1.8330	1.6811
	(0.0084)	(0.0084)	(0.0086)	(0.0102)	(0.0100)	(0.0107)	(0.0102)	[0.1654]
Observations	58,011	70,115	64,111	60,058	64,134	59,692	64,647	–
				Panel d) Top 2%				
Pareto $\hat{\alpha}$	1.6714	1.8614	1.9314	2.0737	2.0976	2.1751	2.2160	2.0038
	(0.0120)	(0.0122)	(0.0132)	(0.0147)	(0.0143)	(0.0154)	(0.0151)	[0.1932]
Observations	38,673	46,742	42,740	40,039	42,756	39,794	43,097	–
				Panel e) Top 1%				
Pareto $\hat{\alpha}$	2.2300	2.4350	2.4470	2.5914	2.5046	2.7075	2.8641	2.5399
	(0.0227)	(0.0225)	(0.0237)	(0.0259)	(0.0242)	(0.0271)	(0.0276)	[0.2052]
Observations	19,337	23,373	21,372	20,020	21,377	19,898	21,551	–

Notes: The table reports the results of OLS rank regressions with $\log\left(\text{rank}(y_i) - 1/2\right) - \log N$ as the dependent variable. Asymptotic standard errors computed as $(2/N)^{1/2}\hat{\alpha}$ are reported in parentheses (see Gabaix and Ibragimov, 2011). The data are a 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level, where each pixel is 30×30 arc seconds. The last column reports the point average of the seven satellites and its standard deviation in brackets.

Table 2 reports the corresponding results. We separately estimate the Pareto shape parameter for each of the radiance-calibrated satellites and for the top 5% to 1% of the data. For our purposes, the most relevant estimates are those for the top 3% to 4%, which corresponds to the share of pixels we will replace later on (see Table H-2). The point estimates in this range are comparatively stable. They only rise steeply once smaller shares are considered.[24] The last column reports the simple average of the estimated coefficients. Somewhat remarkably, the average Pareto shape parameter is about one and a half for the top 4% of lights—not far from our model-based guesstimate.[25] Moreover,

---

[24]In theory, with Pareto-distributed data, conducting the estimation with the portion of the distribution above a higher threshold $y_c^h > y_c$ should lead to the same estimated α. In practice, this will often not hold exactly and the results for tail regressions are known to depend on the precise threshold used (see for instance Rosen and Resnick, 1980).

[25]Our correction uses a fixed Pareto parameter. If we take the results in Table 2 at face value and assume that the true parameter varies across years, then fixing the parameter at the average will lead to an underestimation of the brightness of top-coded pixels in some years and overestimation in others. However, we have no means of differentiating the stability of the true parameter from noise in the

averaging all 21 coefficients in panels (a) to (c) also yields a central estimate of about one and a half. We find only limited variation over time at the lower thresholds, apart from a jump in the early 2000s. Given the noise in the radiance-calibrated data, we are reluctant to interpret these changes as an indication that the global distribution of light intensities (or bright cities) is shifting over time and lack additional data to investigate this possibility. Our range of parameter estimates imply a level of inequality comparable to the top tail of the U.S. income or wealth distribution (Piketty, 2003, Atkinson et al., 2011). Top lights are also considerably more equally distributed than city sizes, just as predicted by the model in the previous section.

We provide an array of additional robustness checks in Online Appendix E and F. Our main findings are robust to *i)* using the Hill estimator instead of OLS rank regressions, *ii)* estimating unrestricted rank regressions, which are another way of testing for a Pareto distribution, and *iii)* using top-coding free data from the new VIIRS satellites, whose first available year is 2015. The VIIRS satellites are technologically superior and recorded at a much finer resolution. Reassuringly, they are as indicative of a Pareto tail in top lights as the radiance-calibrated data and provide even stronger evidence in its favor, such as nearly linear Zipf plots.

4   Correcting for top-coding

We propose a simple correction procedure inspired by the methods used in the top incomes literature (Piketty, 2003, Atkinson et al., 2011). All pixels with light below the top-coding threshold are unaltered, while those above are replaced by a Pareto-distributed counterpart. The replacement uses a simple algorithm which infers the ranks for the theoretical quantiles from the available radiance-calibrated data. An appealing feature of this approach is that it keeps the overwhelming majority of the data intact and replaces only a small but highly influential fraction of pixels. Moreover, the result is a new series of images that can be used flexibly in applications.[26]

---

radiance-calibrated data. Moreover, we would be overstating the brightness of the top tail somewhat in earlier years (e.g., 1.35 rather than 1.5 in 1996) and understating it in later years (1.59 rather than 1.5 in 2010). Using a fixed value is conservative because it biases the growth rate of top-coded pixels downward.

[26]There have been a number of attempts to correct the stable lights data in the remote sensing literature. One type of corrections is based on cubic pixel-level regressions (see e.g. Letu et al., 2012), which is a heavily parametric approach and usually assumes that light emissions were constant over some area or period. An alternative approach uses land cover and vegetation data to non-linearly transform night lights. Zhang et al. (2013), for example, define an index which combines night lights with a Normalized Difference Vegetation Index (NDVI). While this recovers some of the city-wide variation in light, the conflation of lights with NDVI data is not ideal for a number of reasons: *i)* the original scale of the stable lights data is lost, *ii)* all values—not just the top-coded data points—are being adjusted, and *iii)* researchers may wish to use the NVDI and other land cover data as explanatory or dependent variables and could reasonably worry about mechanical correlations introduced by this approach. All of these studies agree that the radiance-calibration method of Hsu et al. (2015) is the relevant benchmark,

Our theoretical arguments and empirical tests suggest a Pareto parameter around *one and a half*. Recall that a plausible parameterization of our model also implies $\alpha = 1/0.7 \approx 1.43$ and our empirical estimates inferred from the radiance-calibrated data are centered on one and a half. We use this fixed value as a rule-of-thumb parameter. Of course, our procedure is not predicated on a particular value, nor does it require the parameter to be constant over time. However, we do not detect an unambiguous trend indicating that the distribution of top lights has become more equal over time. As will become clear shortly, assuming a constant parameter value still allows for significant variation over time as particular pixels *i)* cross the top-coding threshold and *ii)* achieve a higher rank vis-à-vis other pixels. Cities can thus become brighter in absolute terms and grow relative to other cities after the correction.[27]

A remaining challenge is to geo-reference the Pareto quantiles so that the brightest pixels actually end up in the centers of dense urban agglomerations.[28] We propose a straightforward solution. Since we know the exact location of all pixels, we can rank them according to their radiance-calibrated values from the nearest year and distribute the highest values from the Pareto distribution in the same manner. Working with ranks avoids importing the artificial variability of the radiance-calibrated satellites but preserves a crucial part of the data structure.[29] Our algorithm for replacing top-coded pixels with their quantile counterpart from Pareto distribution works as follows:

(1) For each of the 34 satellite-years t of stable lights data, calculate the number N_t of pixels ≥ 55 DN to be replaced.

(2) Produce a ranking of these N_t pixels based on the radiance-calibrated data associated with the same satellite-year or the data from the closest year.[30]

(3) Generate N_t quantiles from a truncated Pareto distribution with rule-of-thumb $\alpha = 1.5$, top-coding threshold $y_c = 55$, and upper bound $H = 2000$.

---

but None of them directly combines the two data sources.

[27]In additional analyses, we assign the seven satellite-specific estimates to the adjacent stable light satellite years. This approach gives very similar results overall, albeit with more jumps. Country-specific estimates also lead to similar results, although comparatively darker and poorer countries experience larger corrections. By using the same parameter for the whole world, we deliberately remain agnostic. Additional results are available on request.

[28]Another option is to analytically correct the relevant summary statistics, such as mean lights and spatial Gini coefficients. We provide closed-form solutions that are particularly useful for back-of-the-envelope calculations in Online Appendix G. Exclusively correcting the summary statistics limits the potential applications and is not sufficient when the actual locations of the top-coded pixels matter.

[29]The ranks of the pixels are much more stable over time than the values of the radiance-calibrated data. The rank correlation of maximum city lights typically ranges from 0.90–0.95 for adjacent radiance-calibrated satellites (see Online Appendix A).

[30]Whenever the radiance-calibrated data create ties, we first try to break those ties using the ranks of the stable lights data and then use the ranks of neighboring radiance-calibrated satellites. This produces a near-unique ranking each time. We also experimented with other rankings. The results are very similar and available on request.

(4) Replace the N_t stable lights pixels ≥ 55 so that the stable lights pixel with the i-th highest rank from (2) is replaced by the i-th highest quantile from (3).

In other words, our procedure combines the desirable features of the stable lights (annual availability and measurement for all non-top-coded values) with those of the radiance calibrated lights (distributional shape of the top as well as stability of the ranks). We apply this procedure to every stable lights satellite image over the entire period from 1992 to 2013. This is the data that we use in the subsequent application and which underlies Figure 1. Note that the support of the Pareto distribution is unbounded, so that its highest quantiles yield a small number of values far exceeding the natural limit of man-made light intensity. To generate a realistic depiction of city lights we impose an upper bound of 2000 DN, which approximately corresponds to the average maximum observed in the world's brightest cities (see Online Appendix A).[31] Using the truncated Pareto barely affects the results in the aggregate but only a few extreme values.[32]

The correction significantly impacts our understanding of the global distribution of economic activity. On average, 3.7% pixels above the top-coding threshold account for 17.7% of the total sum of lights observed in the stable lights data. This share almost doubles to 31.94% after the correction—a statement which is approximately true in every individual satellite-year. However, later shares are bigger in absolute terms before and after the correction. In other words, about four percent of all lit pixels account for about a third of all visible economic activity (if the relationship between light intensity and GDP is linear). The global Gini coefficient rises by nine percentage points, on average, after the correction. Note that the annual variation in the Gini coefficient is only a few percentage points so that the size of the top-coding correction swamps it. These findings are not particularly sensitive to the choice of the Pareto shape parameter, although the size of the top-coding correction varies somewhat. Online Appendix H provides a variety of summary statistics and sensitivity checks using different Pareto shape parameters. The size of the correction in both the country averages of light intensities and the country-wide spatial Gini coefficients varies systematically with GDP per capita, country size, and population density, in line with the occurrence of top-coding illustrated in Figure 3.

Figure 8 illustrates that the correction almost doubles the correlation between the values of the top-coded pixels and their corresponding values in the radiance-calibrated

---

[31]It is difficult to provide a quantitative rationale that links this number to a physical quantity because the radiance-calibrated data are not radiometrically calibrated across years. Hsu et al. (2015, p. 1872) caution explicitly that the "product is only suitable for analyses which do not require actual radiance".

[32]Online Appendix A reports and discusses the observed city-wide maxima obtained by the radiance-calibrated satellites which motivate this upper bound in man-made luminosity. The truncated Pareto has the CDF $F(y) = \left(1 - \left(\frac{y_c}{y}\right)^{\alpha}\right) / \left(1 - \left(\frac{y_c}{H}\right)^{\alpha}\right)$ for $y_c \leq y \leq H$, where an upper bound of $H = 2000$ does not affect the overwhelming majority of stable lights pixels ≥ 55 to be replaced, but does ensure that the values at the very top, say, 0.01% of the data, do not exceed the upper bound in man-made luminosity (or the values typically observed in the radiance-calibrated data). Our results in the subsequent sections are very similar, no matter if we use the simple Pareto or the truncated Pareto.

data. For all pixels above the top-coding threshold, annual correlations between the stable lights and the radiance-calibrated data is only around 0.49–0.53, but jumps to 0.94–0.97 after the correction. Considering that we only use the ranks of the top tail from the radiance-calibrated data together with a rule-of-thumb Pareto fit, the tight pixel-level fit between the corrected and benchmark data is remarkable.

Figure 8 – Correlations with radiance-calibrated data
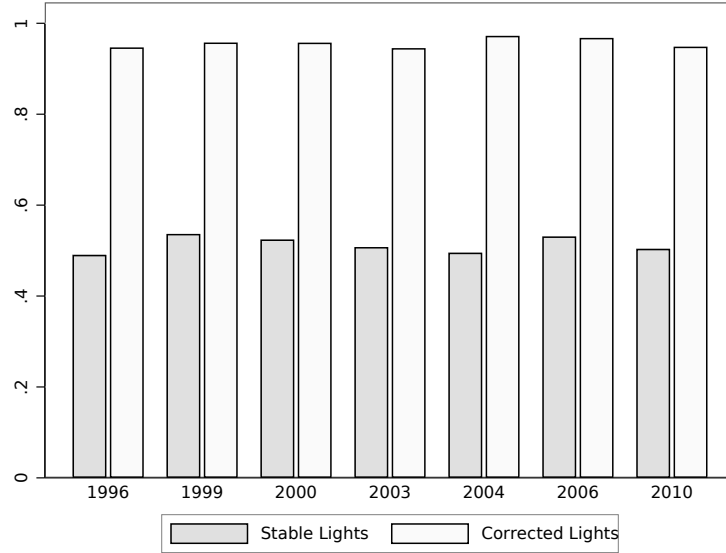
Notes: The figure represents the correlations between the radiance-calibrated lights and, respectively, the uncorrected and corrected lights above DN = 55 for the seven years where the radiance-calibrated data are available. The corrected and uncorrected data are based on satellites F12 in 1996, F14 in 1999, F15 in 2000, F15 in 2003, F16 in 2004, F16 in 2006, and F18 in 2010.

We carry out three additional benchmarking exercises to assess the properties of the new data and better understand at which scale of aggregation top-coding influences the conclusions we are likely to draw from nighttime lights. We only briefly summarize the results here and relegate a full discussion to Online Appendix I. The first exercise estimates national light-output elasticities (as in Henderson et al., 2012) and investigates non-linearity in the light production function (as in Hu and Yao, 2019). Even at this high level of aggregation, the top-coding corrected series performs slightly better (in the sense of marginally higher elasticities, higher R^2s, and less non-linearity). For the most part, however, the estimates using the corrected data are not just economically but also numerically very close to the original results, no matter which way around we study the relationship. Our second benchmark goes from the national to the regional level and investigates non-linearity in the light production function using GDP data for German districts and U.S. counties. Focusing on two developed economies with high-quality regional accounts allows us to directly study the structural relationship (of lights on GDP), as measurement errors in GDP are likely to be small and unrelated to measurement errors in nighttime lights. We obtain two notable results. First, the correction removes a

substantial portion of non-linearity of the income elasticity of lights in the time series and cross-section. Both the German and U.S. data suggest that the corrected lights scale more linearly in economic density than the top-coded stable lights. Second, light intensities in the corrected series are approximately linear in population density, whereas the stable lights respond less to income differences in densely populated areas. Both findings suggest that the corrected data better capture the light-output gradient in developed economies with mature urban structures. We expect that this pattern is not limited to developed economies, provided that the regions under investigation are small enough to include a significant number of top-coded pixels. To test this claim, we investigate the relationship between average household wealth and light intensity around more than 7,000 urban sample clusters from the Demographic and Health Surveys (DHS) undertaken in 29 African countries from 1992 to 2013. We find some evidence of non-linearity in the relationship of wealth to the stable lights data which is either weakened substantially or disappears when the corrected data are used.

5   Application: Cities in Sub-Saharan Africa

Armed with this new data, we now return to the question of whether primate cities are outgrowing secondary cities in Sub-Saharan Africa. In most African countries, economic activity in terms of both population and value-added is concentrated in the biggest city. Dar es Salaam, Kinshasa, and Lagos already are megacities with more than 10 million inhabitants or will soon attain that status (United Nations, 2018). Strong spatial concentration is often a feature of countries with poor infrastructure and a low level of development (Krugman, 1991, Puga, 1998, Jedwab et al., 2019), although Africa's urbanization differs from the historical experience of industrialized countries.

The new millennium marked a turning point for most African economies. Sustained consumption growth and pro-poor distributional change brought about falling poverty rates (Bluhm et al., 2018). However, as countries are developing, it is theoretically unclear whether secondary cities will catch up (Duranton, 2008). There is some empirical evidence suggesting that this is the case. Henderson et al. (2012), for example, estimate that the African hinterland was growing about 2.3% faster than primate cities over the period from 1992 to 2008.[33] Moreover, many in the World Bank view secondary city development as key to sustained poverty reduction (Christiaensen and Todo, 2014, Christiaensen and Kanbur, 2017). Yet, manufacturing is heavily concentrated in primate and coastal cities with greater access to world markets. The oil price boom of the 2000s, for example, hurt remote secondary cities more than primate cities (Storeygard, 2016). Secondary cities in

---

[33]Note that Henderson et al. (2012) also suggest that a "detailed study would be required to explain the result" (p. 1024). In India, Gibson et al. (2017) find that secondary cities matter more for rural poverty reduction when studying the spillover of urban growth at the intensive and extensive margin.

Africa have also been characterized as "consumption cities", catering to the agricultural hinterland rather than the modern sector (Gollin et al., 2016).

Whether or not primate cities are driving productivity growth depends—at least in part—on their internal structure. Informal settlements in large cities can more easily absorb growing populations (Jedwab et al., 2019), but if these neighborhoods remain badly connected to the center, then cities will be fragmented and less productive (Lall et al., 2017). An adverse urban form implies that companies and inhabitants are faced with high transport costs and long commuting times, limiting interactions and positive spillovers across the city (Rosenthal and Strange, 2004, Harari, 2020, Venables, 2017).

Our application tackles both of these aspects. In the remainder, we use the top-coding corrected data to show that primate cities have grown at least as quickly as secondary cities and towns over the period from 1992 to 2013. Online Appendix J adds an explorative study of fragmentation in African cities. We find that inequality in light decreased over the sample period, in particular in primary cities, while there is also evidence of increased fragmentation.

City boundaries: Urban areas are often delineated using night light by defining them as contiguously lit clusters above some fixed luminosity threshold (Small et al., 2011, Storeygard, 2016, Harari, 2020). Satellite-derived footprints are particularly useful in Africa, where administrative boundaries quickly become outdated. Relying on nighttime lights is appealing, since they offer a time-varying measure of urban expansion, but the thresholding approach suffers from a well-known problem: no single threshold works well for all cities. Thresholding overestimates the urban extent of larger cities and penalizes other cities at the same time (Small et al., 2011, Abrahams et al., 2018). We address these issues using a method developed by Abrahams et al. (2018). Their de-blurring algorithm reduces the non-linear "overglow" in the lights data and improves the accuracy of the identified urban extents.[34]

To capture changes at the extensive margin and minimize measurement errors, we define cities as contiguously illuminated pixels in the de-blurred lights, provided that a light source is detected in at least two satellites over a period of three years.[35] The period from 1992 to 1994 marks the initial boundaries and the period from 2011 to

---

[34]The de-blurring approach is based on two insights into the data generating process: *i)* the original light sources are blurred by a symmetric Gaussian point-spread function, and *ii)*, pixels in which light sources are located must be local maxima in the so-called percent frequency of light detection image (for more detail see Abrahams et al., 2018). Note that we have used light intensity thresholds for the detection of city boundaries in a previous version of this paper which yields qualitatively similar results. We do not use their de-blurring procedure as part of the top-coding correction (to retain a complete ranking of pixels). However, the corrected data can be de-blurred ex-post just like the original stable lights product.

[35]Two images of the same place in the same year may indicate different urban footprints due to a lack of onboard calibration. Considering a window of three years and requiring multiple detection points effectively cancels out most of this artificial variation.

Figure 9 – Primate and secondary cities in Sub-Saharan Africa
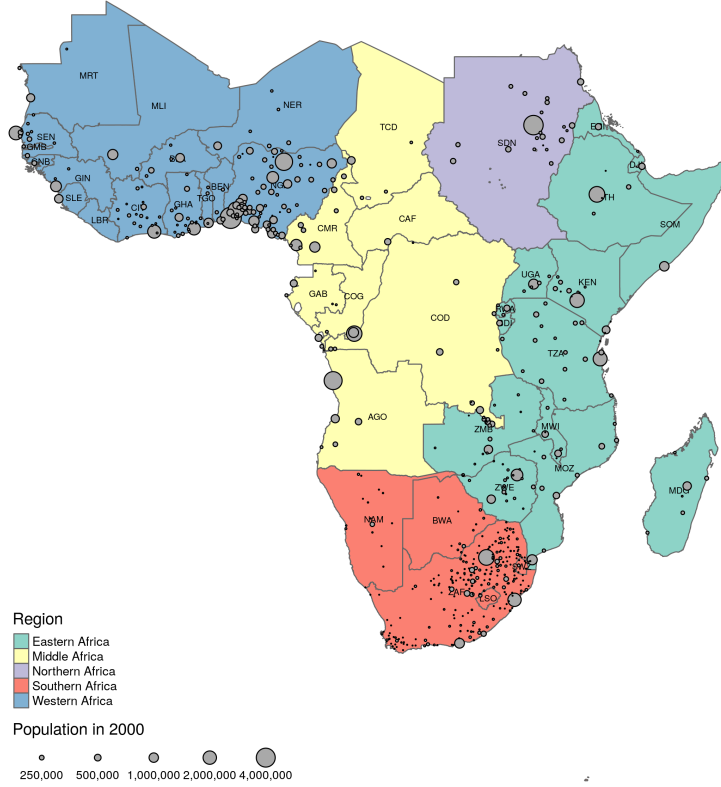
Notes: Illustration of the location of our panel of 41 primate and 521 secondary cities in SSA using the urban footprint detection algorithm outlined in the text. Note that we dropped Equatorial Guinea due to gas flaring on the capital island of Malabo and consider South Sudan as part of Sudan for the entire sample.

2013 represents the latest available boundaries.[36] We then identify city locations by overlaying these urban areas with all settlement points from the Global Rural-Urban Mapping Project (GRUMP) within three kilometers of the urban extent.[37] An urban area receives the name and attributes of the most populous settlement located within the expanded urban extent.

We define three different margins based on these urban areas. First, we track how the economic activity of a city develops over time at the intensive margin using only the initial boundaries of a city. Second, we take the union of the initial and final period boundaries to identify the "envelope" or maximum urban extent of the cities in our sample. Third, we separate the intensive margin from extensive growth in the fringe by taking the difference of the envelope to the initial data. Several smaller cities merge over

---

[36]Online Appendix J shows both the initial and latest boundaries together with daytime images of Lagos, Luanda, and Johannesburg taken at the end of both periods.

[37]We manually extend this data to include all coordinates of cities which at some point over the period from 1992 to 2013 were designated the administrative capital of a province or state. We first identify the administrative capitals of subnational regions using www.statoids.com and then geocode each city using multiple online gazetteers.

time or become absorbed by a larger city in their vicinity. We aggregate all parts of a city that will eventually become a single agglomeration to deal with this issue. We then limit our analysis to agglomerations observed in the initial and final periods. Finally, we identify the *primate city* in each country as the city with the largest population and define all other cities and towns as secondary. Not all small settlements qualify as secondary cities but there is no consensus on their definition (minimum sizes typically range from 100,000 to 500,000 people). Since we lack accurate population data for many smaller settlements, we work with an area threshold of 9 km²—about 3 × 3 pixels—and show that our results are robust to variations of this threshold in Online Appendix J.[38]

Figure 9 shows a map of our universe of 562 cities and their population in 2000 (if available). 41 cities, one per country, are primate cities, while 521 are secondary cities. The five largest urban agglomerations in the sample (according to their initial area) are Johannesburg, Durban, Cape Town, Lagos, and Harare.

City growth: Cities grow at the intensive margin, as existing city quarters develop and become brighter, and at the extensive margin, as they expand and new lights are added to the surrounding areas. Total city growth is the combination of these two margins. We examine total city growth and its components separately.[39] This allows us to gain insights into *where* cities are growing, and which components of city growth are (more severely) mismeasured due to top-coding.

Table 3 compares the sum of lights for the commonly used stable lights data and our corrected counterparts in 1992 and 2013. Panel (a) shows results for the envelopes (i.e. the maximum extent) and panel (b) for the initial boundaries. No matter which boundaries we use, the sum of light of the average primary city is more than ten times as large as that of the average secondary city.

Substantial differences appear once we focus on city growth or compare the two groups towards the end of the sample. The raw data shows that *i)* top-coding affects the average primary city much more than the average secondary city, and *ii)* the size of the correction becomes substantially larger over time. The correction adds 11.1% to the average sum of lights of primate cities in 1992 but increases it by 27.4% in 2013. This has implications for our understanding of agglomeration economies and city growth in Sub-Saharan Africa. The stable lights data show that secondary cities are on average growing

---

[38]We do not use a population threshold since we do not have population estimates for many of the smaller cities. Noise in the measurement of the smaller settlements does not drive our overall results. On the contrary, the difference between primate and secondary cities increases at higher pixel thresholds (see Table J-3 in Online Appendix J).

[39]The decomposition we have in mind is $\Delta \ln DN_t^{Total} = \omega \Delta \ln DN_t^{Core} + (1 - \omega)\Delta \ln DN_t^{Fringe}$, where ω is the area of the core relative to the area of the envelope (in the final period). Hence, total growth in primary or secondary cities is a weighted average of the growth rates of lights in the core (the intensive margin) and growth rates of lights in the fringe (the extensive margin). All of these quantities, including the ωs, vary at the city-level, so that the decomposition cannot be applied directly to the summary statistics or regression results presented in this section.

Table 3 – Summary statistics for African cities, 1992–2013, envelopes and initial boundaries

	Stable Lights		Corrected Lights	
	Primate	Secondary	Primate	Secondary
	Panel a) Sum of lights in the envelopes			
Sum in 1992	22110.12	1970.52	24565.71	2075.82
	(72342.16)	(6682.16)	(83637.07)	(7653.71)
Sum in 2013	36242.39	3166.52	46172.05	3494.81
	(94385.49)	(8801.42)	(128885.78)	(11575.72)
Av. Growth	0.0235	0.0226	0.0300	0.0248
	Panel b) Sum of lights in initial boundaries			
Sum in 1992	20321.93	1778.94	22777.48	1884.24
	(69367.84)	(6489.79)	(80659.66)	(7471.89)
Sum in 2013	27485.29	2472.93	37126.00	2798.45
	(85385.54)	(8122.60)	(119247.35)	(10975.28)
Annualized growth rate	0.0144	0.0157	0.0233	0.0188

Notes: The table reports a selection of summary statistics for African cities based on their envelopes as maximal extents in panel (a) and their initial boundaries in panel (b). Standard deviations are given in parentheses. Annualized growth rates are computed as $\frac{1}{21}(\ln x_{2013} - \ln x_{1992})$, where x refers to the data per group reported in the table.

at the approximately same rate as primary cities—an artificial byproduct of more and more pixels in primate cities becoming top-coded. The top-coding corrected data show that primary cities on average outgrew secondary cities by about half a percentage point. Panel (b) shows that these differences in total city growth are mainly driven by growth at the intensive margin. Dense city centers are typically the places where top-coding is most pronounced. Online Appendix J adds descriptive data for each country. It confirms that top-coding plays virtually no role in small primate cities, such as Bissau, while the annualized growth rates of larger cities, like Kinshasa, Johannesburg, Luanda, Dakar, or Khartoum, increase by about a percentage point or more. Nevertheless, it holds across countries that the top-coding correction tends to affect primary cities much more strongly than secondary ones.

Figure 10 adds another piece of evidence along these lines. We now use the difference between the two data sources on the left-hand side of a fixed-effects regression with year-specific coefficients on the primacy status of a city. The coefficient plot shows that the difference in total light emitted (in the city envelopes) between primate and secondary cities is considerably larger in the corrected data than the top-coded data in all years. This difference, which can be fully attributed to top-coding, starts to rise strongly after 2005. As a result, both cross-sectional differences and relative growth rates are increasingly distorted in the later years of the stable lights data. Results using the initial boundaries are virtually identical (not reported).

We underpin these descriptive results with panel growth regression for the intensive

Figure 10 – Difference between primate and secondary cities in the two data sources
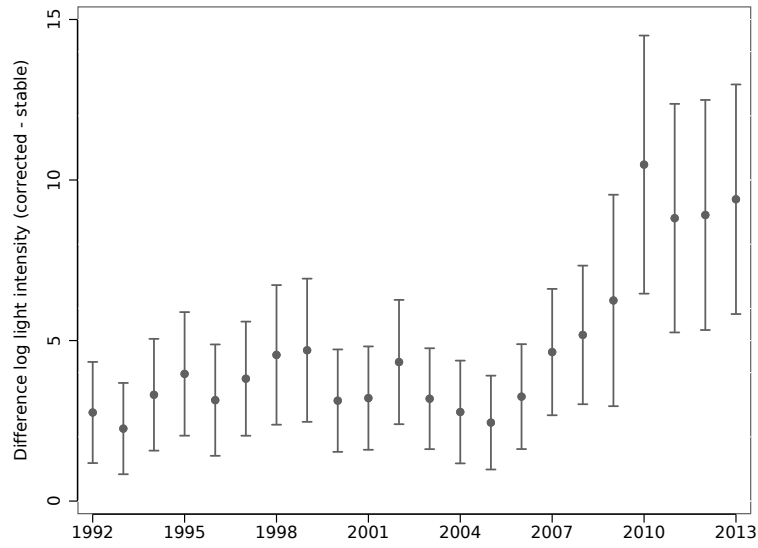
Notes: Illustration of how the component of the cross-sectional difference between primate and secondary cities which can be fully attributed to the top-coding correction evolves over time. The figure plots the coefficients, β_p, and the corresponding 95% confidence intervals obtained from the following regression: $\ln \Delta_{ijt} = \sum_{p=1}^{T} \beta_p (s_p \times P_{ij}) + s_t + \epsilon_{ijt}$ where $\ln \Delta_{ijt}$ is the log difference between the corrected data and the stable lights data in city i within country j at time t computed using the union of final and initial boundaries of a city (the envelopes), P_{ij} is an indicator for primate cities, s_p are indicators that are one when $p = t$ and zero otherwise, and s_t are year fixed effects. Standard errors are clustered at the city level.

Table 4 – Growth regressions for African cities: intensive, extensive and total growth

	Stable lights		Corrected lights		Radcal lights	
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Varies by panel					
	Panel a) Intensive margin: log lights in the initial footprint					
Primate × trend	0.215	-0.064	0.677	0.514	1.047	0.594
	(0.183)	(0.266)	(0.254)***	(0.228)**	(0.397)***	(0.230)***
	[0.193]	[0.269]	[0.256]***	[0.228]**	[0.403]***	[0.280]**
Observations	12364	12364	12364	12364	3934	3934
Cities	562	562	562	562	562	562
	Panel b) Extensive margin: log lights in the fringe					
Primate × trend	2.955	2.830	2.992	2.883	2.553	2.207
	(0.433)***	(0.304)***	(0.442)***	(0.309)***	(0.625)***	(0.449)***
	[0.433]***	[0.303]***	[0.442]***	[0.307]***	[0.621]***	[0.426]***
Observations	11410	11410	11410	11410	3633	3633
Cities	519	519	519	519	519	519
	Panel c) Total growth: log lights in the envelope					
Primate × trend	0.888	0.540	1.168	0.902	1.485	0.971
	(0.299)***	(0.265)**	(0.332)***	(0.245)***	(0.445)***	(0.241)***
	[0.300]***	[0.265]**	[0.331]***	[0.243]***	[0.446]***	[0.268]***
Observations	12364	12364	12364	12364	3934	3934
Cities	562	562	562	562	562	562
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	–	✓	–	✓	–
Country-Year FE	–	✓	–	✓	–	✓

Notes: The table reports the results of city-level panel regressions using the stable lights, the top-coding corrected, and radiance-calibrated data. Panel (a) considers growth within the initial boundaries of a city, panel (b) shows results for the fringe, defined as the difference between the envelope of a city and the initial boundary, and panel (c) show the results for the envelope of a city. Panel (b) reports fewer observations because the urban extent of some cities does not expand during our period of study, so that the fringe is undefined. Areas with zero or few lights are adjusted following the procedure in Storeygard (2016) where the minimal detectable light intensity (in DN) for a given area is added to the total before dividing by the area of the initial boundary, fringe, or envelope. All coefficients are scaled by 100 for readability. The specifications are variants of $\ln \text{LIGHTS}_{ijt} = \beta_1 t + \beta_2(t \times P_{ij}) + c_{ij} + s_{jt} + \epsilon_{ijt}$ where t is a linear time trend, P_{ij} is an indicator for primate cities, c_{ij} is a city fixed effect and s_{jt} contains a varying set of fixed effects (year, or country-year). Standard errors clustered at the city level are reported in parentheses. Conley errors with a spatial cutoff of 1,000 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

margin, extensive margin, and total envelope. In each case, we regress the log of lights per km^2 recorded for city i in country j at time t, or $\ln \text{LIGHTS}_{ijt}$, on an interaction of the linear time trend, t, with an indicator for primate cities, P_{ij}, city fixed effects, c_{ij}, and a set of year or country-year fixed effects, s_{jt}. The areas of the core, fringe, or envelope are fixed over the entire sample (and defined by the initial and final boundaries of a city). Annual fluctuations in the light sensitivity of the satellites imply that the

baseline growth rate (the coefficient on t without the interaction with primacy) cannot be identified, which is why we focus on the difference in growth rates and absorb the trend in time fixed effects. The differential between primate and secondary cities can still be estimated in our specifications and is net of measurement error which is common to all cities (in a country) but varies by year.

Table 4 confirms that primate cities outgrew secondary cities by a significant margin in the corrected data. Columns (1) and (2) in panel (a) show that the stable lights data suggest no significant differences in intensive growth across city types. On the contrary, the corrected data in columns (3) and (4) indicate that primate cities grew 0.51-0.68 percentage points faster per year than secondary cities. The estimated coefficient on the interaction term directly corresponds to a difference-in-means test. Hence, we can reject the null hypothesis that both trends are the same at conventional significance levels. Columns (5) and (6) add that we find similar trends in the radiance-calibrated data for the seven years in which they are available. In fact, the results are very similar once country-year fixed effects are included. If we take these estimates at face value, then average light output in primary cities was 10.7-14.3% larger than in secondary cities at the end of the 21-year period.

So far we have focused exclusively on the intensive margin. When the city footprint is kept fixed, we can interpret growth in lights directly as increases in population density and economic activity per square kilometer. Population density, in turn, is strongly correlated with living standards and public good provision in developing countries (Gollin et al., 2017). Dense city centers are also the places where top-coding is most pronounced.

Panel (b) in Table 4 shows the estimated differential growth rates at the fringe (i.e., the areas into which the city expands after the initial period). New developments at the fringe of a city are usually dimmer. Urban sprawl is typically suburban or peri-urban, and density only increases later when these areas become part of the city proper. However, primate cities have been growing substantially faster than secondary cities, with a differential of 2–3 percentage points per annum, according to all three data sources. As expected, the top-coding correction does not have a noticeable impact on the urban fringe. The coefficients on the stable lights and corrected data are numerically very close. In contrast, the radiance calibrated lights, which are less reliable at low light intensities, somewhat understate growth in the fringe.

Finally, panel (c) shows that total growth in primate cities is also understated by the stable lights data but less than growth at the intensive margin. In the strictest specification with country-year fixed effects, we estimate a differential of 0.9 percentage points per annum using the corrected data instead of only 0.54 percentage points using the stable lights data. All corrected estimates lie between the stable lights and radiance calibrated lights. They are closer to the latter because the core represents, on average, about 71.7% of the total envelope of all cities.

Table 5 – Population growth regressions for African cities, Africapolis data

	Dependent variable: Log Population									
	All Cities		Cities ≥ 10,000 pop in 1990		Cities ≥ 25,000 pop in 1990		Cities ≥ 50,000 pop in 1990		Cities ≥ 75,000 pop in 1990	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Primate × trend	0.596	0.387	0.602	0.328	0.585	0.382	0.617	0.401	0.671	0.497
	(0.267)**	(0.210)*	(0.268)**	(0.217)	(0.278)**	(0.248)	(0.280)**	(0.248)	(0.299)**	(0.247)**
	[0.279]**	[0.212]*	[0.279]**	[0.220]	[0.293]*	[0.249]	[0.313]**	[0.249]	[0.309]**	[0.226]**
City FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	–	✓	–	✓	–	✓	–	✓	–
Country-Year FE	–	✓	–	✓	–	✓	–	✓	–	✓
Observations	13662	13662	7198	7198	2955	2955	1412	1412	939	939
Cities	4513	4513	1848	1848	759	759	361	361	239	239

Notes: The table reports the results of city-level panel regressions using the population data from Africapolis. Three different samples of SSA cities are selected by setting varying lower thresholds for the population in 1990. The sample years are 1990, 2000, 2010 and 2015. The specifications are variants of $\ln \mathrm{POP}_{ijt} = \beta_1 t + \beta_2 (t \times P_{ij}) + c_{ij} + s_{jt} + \epsilon_{ijt}$ where t is a linear time trend, P_{ij} is an indicator for primate cities, c_{ij} is a city fixed effect and s_{jt} contains a varying set of fixed effects (year, or country-year). All coefficients have been scaled by 100 for readability. Standard errors clustered at the city level are reported in parentheses. Conley errors with a spatial cutoff of 1,000 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Taken together, Table 4 shows how our correction joins the proprieties of the radiance-calibrated data where they are superior (in the urban centers) with those of the stable lights data where they are more reliable (at the fringe). Online Appendix J presents two robustness checks for these findings. Table J-3 and Table J-4 show that the growth differential remains intact when we vary the area threshold for the minimum secondary city size or omit any one region of Sub-Saharan Africa.

Finally, we test if the pattern of relatively faster primate city growth is an artifact (of some versions) of the lights data or if it matches relative trends in population growth. While censuses are typically undertaken at a much lower frequency and are subject to a variety of measurement errors, their data generating process is independent of that of nighttime lights. Moreover, our preferred interpretation of our main results is that differences in population density at the city core are driving overall differences in relative growth rates, implying that we should see similar trends in population data. The Africapolis database (OECD/SWAC, 2020) compiles local census area estimates, identifies the urban extent of agglomerations using built-up data, groups localities together if they are part of the same agglomeration, and then interpolates the census-derived population estimates to reference years (decades since 1950). Nighttime lights are not used at any step in this process. We use the Afripolis data in 1990, 2000, 2010, and 2015 to approximately match our sample period, and the same countries as in our main city sample (apart from Madagascar, which is not available in Africapolis). Note that our approach and Africapolis use different definitions of agglomerations so that there is no one-to-one correspondence between the two samples at the city level.

Table 5 shows that the relative trends in population growth are similar to those obtained via the corrected lights. Columns (1) and (2) mirror the specifications from Table 4 for all cities in the Africapolis data; the remaining columns then repeat these two specifications using progressively larger initial population thresholds. We find a large and significant growth differential in all specifications with country and year fixed effects, as well as the full city sample and the sample of cities with more than 75,000 inhabitants when country-year fixed effects are included. In all other samples, the differential remains in between these two estimates but the t-statistics fall below conventional thresholds. Given the limited time variation, we take this as sufficient evidence to suggest that the differences in light intensities between primary and secondary cities over time are primarily capturing differences in population growth rates.

6 Concluding remarks

While satellite data of nighttime lights are a popular proxy for economic activity, they suffer from top-coding and severely underestimate the brightness of most cities. The key contribution of this paper is to provide a solution to this problem and establish new

findings regarding the economic performance of cities in Sub-Saharan Africa.

Our solution rests on the claim that top lights can be characterized by a Pareto distribution. We support this conjecture in two ways. First, a model of luminosity emitted by large cities suggests that plausible assumptions directly lead to a power law in light emissions. Second, a battery of empirical tests indicates that a Pareto distribution is a sound representation of the data. Other parametric or non-parametric approaches are possible, but we find it appealing to directly link the distribution of bright lights to both Zipf's law of cities and the standard tail extrapolation problem. On this basis, we develop a geo-referenced ranking procedure to replace the top-coded pixels with their theoretical counterparts and present a new global panel of light intensities over the period from 1992 to 2013.

The new data lends itself to numerous applications and performs well in several benchmarking exercises. In this paper, we focus on city growth. Our main finding is that primary cities have cemented their dominant position and outpaced growth in secondary cities. We only obtain this result using nightime lights data that is not top-coded and verify these trends using population data. Institutional features certainly play a role in the continued development of primary cities and warrant more research. Finally, by focusing on cities in Sub-Saharan Africa, we submit our data to a conservative test. Given that the top-coding correction makes a substantial difference in a setting where electrification rates and urban building densities are low, it will certainly play a larger role in other parts of the world.

References

Abrahams, A., C. Oram, and N. Lozano-Gracia (2018). Deblurring DMSP nighttime lights: A new method using gaussian filters and frequencies of illumination. *Remote Sensing of Environment 210*, 242–258.

Ades, A. F. and E. L. Glaeser (1995). Trade and circuses: Explaining urban giants. *Quarterly Journal of Economics 110*(1), 195–227.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016). Ethnic inequality. *Journal of Political Economy 124*(2), 428–488.

Atkinson, A. B., T. Piketty, and E. Saez (2011). Top incomes in the long run of history. *Journal of Economic Literature 49*(1), 3–71.

Barrios, S., L. Bertinelli, and E. Strobl (2006). Climatic change and rural–urban migration: The case of Sub-Saharan Africa. *Journal of Urban Economics 60*(3), 357–371.

Batty, M. and P. Longley (1994). *Fractal cities: A geometry of form and function.* San Diego, CA and London: Academic Press.

Bertaud, A. and S. Malpezzi (2014). The spatial distribution of population in 57 world cities: The role of markets, planning, and topography. Unpuplished, University of Wisconsin-Madison.

Bettencourt, L. M. A. (2013). The origins of scaling in cities. *Science 340*(6139), 1438–1441.

Bluhm, R., D. de Crombrugghe, and A. Szirmai (2018). Poverty accounting. *European Economic Review 104*(C), 237–255.

Bluhm, R. and G. C. McCord (2022). What can we learn from nighttime lights for small geographies? measurement errors and heterogeneous elasticities. *Remote Sensing 14*(5).

Bonfatti, R. and S. Poelhekke (2017). From mine to coast: Transport infrastructure and the direction of trade in developing countries. *Journal of Development Economics 127*, 91–108.

Brueckner, J. K. (1982). A note on sufficient conditions for negative exponential population densities. *Journal of Regional Science 22*(3), 353–359.

Castells-Quintana, D. (2017). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics 98*, 158–173.

Castells-Quintana, D., T. McDermott, and M. Krause (2021). The urbanising force of global warming: The role of climate change in the spatial distribution of population. *Journal of Economic Geography 21*(4), 531–556.

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.

Christiaensen, L. and R. Kanbur (2017). Secondary towns and poverty reduction: Refocusing the urbanization agenda. *Annual Review of Resource Economics 9*, 405–419.

Christiaensen, L. and Y. Todo (2014). Poverty reduction during the rural-urban transformation: The role of the missing middle. *World Development 63*(C), 43–58.

Cirillo, P. (2013). Are your data really Pareto distributed? *Physica A: Statistical Mechanics and its Applications 392*(23), 5947–5962.

Coleman, J. S. (1964). *Introduction to Mathematical Sociology.* New York, NY: The Free Press.

Cowell, F. and E. Flachaire (2007). Income distribution and inequality measurement: The problem of extreme values. *Journal of Econometrics 141*, 1044–1072.

Desmet, K. and E. Rossi-Hansberg (2013). Urban accounting and welfare. *American Economic Review 103*(6), 2296–2327.

Donaldson, D. and A. Storeygard (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives 30*(4), 171–198.

Duranton, G. (2008). Viewpoint: From cities to productivity and growth in developing countries. *Canadian Journal of Economics 41*(3), 689–736.

Düben, C. and M. Krause (2021). Population, light, and the size distribution of cities. *Journal of Regional Science 61*(1), 189–211.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review 94*(5), 1429–1451.

Eeckhout, J. (2009). Gibrat's law for (all) cities: Reply. *American Economic Review 99*(4), 1676–1683.

Elvidge, C. D., K. E. Baugh, J. B. Dietz, T. Bland, P. C. Sutton, and H. W. Kroehl (1999). Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sensing of Environment 68*(1), 77–88.

Elvidge, C. D., D. Ziskin, K. E. Baugh, B. T. Tuttle, T. Ghosh, D. W. Pack, E. H. Erwin, and M. Zhizhin

(2009). A fifteen year record of global natural gas flaring derived from satellite data. *Energies 2*(3), 595–622.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics 114*(3), 739–767.

Gabaix, X. and R. Ibragimov (2011). Rank - 1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics 29*(1), 24–39.

Gabaix, X. and Y. Ioannides (2004). The evolution of city size distribution. In J. V. Henderson and J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (1 ed.), Volume 4, Chapter 53, pp. 2341–2378. Elsevier.

Gibrat, R. (1931). *Les inegalités économiques: Applications: Aux inegalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc. D'une loi nouvelle, la loi de l'effet proportionnel.* Paris: Librairie du Recueil Sirey.

Gibson, J., G. Datt, R. Murgai, and M. Ravallion (2017). For India's rural poor, growing towns matter more than growing cities. *World Development 98*, 413–429.

Gibson, J., S. Olivia, and G. Boe-Gibson (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics 149*, 102602.

Glaeser, E. and J. V. Henderson (2017). Urban economics for the developing world: An introduction. *Journal of Urban Economics 98*, 1–5.

Glaeser, E. L. (2014). A world of cities: The causes and consequences of urbanization in poorer countries. *Journal of the European Economic Association 12*(5), 1154–1199.

Gollin, D., R. Jedwab, and D. Vollrath (2016). Urbanization with and without industrialization. *Journal of Economic Growth 21*(1), 35–70.

Gollin, D., M. Kirchberger, and D. Lagakos (2017). In search of a spatial equilibrium in the developing world. Working Paper 23916, National Bureau of Economic Research.

Harari, M. (2020). Cities in bad shape: Urban geometry in India. *American Economic Review 110*(8), 2377–2421.

Henderson, J. V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth 8*(1), 47–71.

Henderson, J. V., T. Squires, A. Storeygard, and D. Weil (2018). The global distribution of economic activity: Nature, history, and the role of trade. *Quarterly Journal of Economics 133*(1), 357–406.

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. *Quarterly Journal of Economics 129*(2), 995–1033.

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

Hu, Y. and J. Yao (2019). Illuminating economic growth. IMF Working Paper no. 19/77, International Monetary Fund.

Jean, N., M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon (2016). Combining satellite imagery and machine learning to predict poverty. *Science 353*(6301), 790–794.

Jedwab, R. and A. Moradi (2016). The permanent effects of transportation revolutions in poor countries: Evidence from Africa. *Review of Economics and Statistics 98*(2), 268–284.

Jedwab, R., D. Pereira, and M. Roberts (2019). Cities of workers, children, or seniors? Age structure and economic growth in a global cross-section of cities. Technical report.

Jones, A. (1975). Density-size rule, a further note. *Urban Studies 12*(2), 225–228.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy 99*(3), 483–499.

Lall, S. V., J. V. Henderson, and A. J. Venables (2017). *Africa's cities: Opening doors to the world.* Washington, DC: World Bank.

Lessmann, C. and A. Seidel (2017). Regional inequality, convergence, and its determinants: A view from outer space. *European Economic Review 92*(B), 110–132.

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

Lipton, M. (1977). *Why poor people stay poor: Urban bias in world development.* Cambridge, MA: Harvard University Press.

Luckstead, J. and S. Devadoss (2014). Do the world's largest cities follow Zipf's and Gibrat's laws?

Economics Letters 125(2), 182–186.

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

Michalopoulos, S. and E. Papaioannou (2014). National institutions and subnational development in Africa. *Quarterly Journal of Economics 129*(1), 151–213.

Michalopoulos, S. and E. Papaioannou (2018). Spatial patterns of development: A meso approach. *Annual Review of Economics 10*(1), 383–410.

Mills, E. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review 57*(2), 197–210.

Montalvo, J. G., M. Reynal-Querol, and J. C. M. Mora (2021). Measuring Inequality from Above. Working Papers 1252, Barcelona Graduate School of Economics.

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics 46*(5), 323–351.

Nordhaus, W. (2012). Economic policy in the face of severe tail events. *Journal of Public Economic Theory 14*(2), 197–219.

OECD/SWAC (2020). *Africa's Urbanisation Dynamics 2020: Africapolis, Mapping a New Urban Geography.* Paris: OECD Publishing. West African Studies.

Parr, J. (1985). A population-density approach to regional spatial structure. *Urban Studies 22*(4), 289–303.

Piketty, T. (2003). Income inequality in France, 1901–1998. *Journal of Political Economy 111*(5), 1004–1042.

Pinkovskiy, M. (2017). Growth discontinuities at borders. *Journal of Economic Growth 22*(2), 145–192.

Pinkovskiy, M. and X. Sala-i Martin (2016). Lights, camera ... income! Illuminating the national accounts-household surveys debate. *Quarterly Journal of Economics 131*(2), 579–631.

Pinkovskiy, M. and X. Sala-i Martin (2020). Shining a light on purchasing power parities. *American Economic Journal: Macroeconomics 12*(4), 71–108.

Puga, D. (1998). Urbanization patterns: European versus less developed countries. *Journal of Regional Science 38*(2), 231–252.

Rosen, K. T. and M. Resnick (1980). The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics 8*(2), 165–186.

Rosenthal, S. and W. Strange (2004). Evidence on the nature and sources of agglomeration economics. In J. Henderson and J. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2119–2171. Elsevier North Holland.

Rozenfeld, H., D. Rybski, X. Gabaix, and H. Makse (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review 101*(5), 2205–2225.

Small, C., C. Elvidge, D. Balk, and M. Montgomery (2011). Spatial scaling of stable night lights. *Remote Sensing of Environment 115*(2), 269–280.

Smeed, R. (1961). *The traffic problem in towns.* Manchester Statistical Society Papers.

Soo, K. T. (2005). Zipf's law for cities: A cross-country investigation. *Regional Science and Urban Economics 35*(3), 239–263.

Stewart, J. Q. (1947). Suggested principles of "social physics". *Science 106*(2748), 179–180.

Storeygard, A. (2016). Farther on down the road: Transport costs, trade and urban growth in Sub-Saharan Africa. *Review of Economic Studies 83*(3), 1263–1295.

Tuttle, B., S. Anderson, P. Sutton, C. Elvidge, and K. Baugh (2013). It used to be dark here. *Photogrammetric Engineering and Remote Sensing 79*(3), 287–297.

United Nations (2018). *World urban prospects, the 2018 revision.* United Nations New York.

Venables, A. J. (2017). Breaking into tradables: Urban form and urban function in a developing city. *Journal of Urban Economics 98*, 88–97.

Zhang, Q., C. Schaaf, and K. C. Seto (2013). The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity. *Remote Sensing of Environment 129*, 32–41.

Zielinski, K. (1980). The modelling of urban population density: A survey. *Environment and Planning A 12*(2), 135–154.

Ziskin, D., K. Baugh, F. C. Hsu, T. Ghosh, and C. Elvidge (2010). Methods used for the 2006 radiance lights. *Proceedings of the Asia-Pacific Advanced Network 30*, 131–142.

Online appendix

A   Additional summary statistics

In this section, we provide further details on the construction of the stable lights or radiance-calibrated data and compare their characteristics.

The main advantage of the stable lights series is that they are available as an annual panel from 1992 until 2013. Moreover, for several years more than one satellite orbited Earth, resulting in a total of 34 satellite years. The radiance-calibrated data, by contrast, are based on rare flights of satellites that were about to be decommissioned and could be operated with a different gain setting (lower or higher amplification settings). These auxiliary data are only available for seven years over the entire period from 1996 to 2010. NOAA blends the stable lights data from normal flight operations with these data to obtain the radiance-calibrated series (Elvidge et al., 1999, Ziskin et al., 2010, Hsu et al., 2015). The resulting night light intensities are free of top-coding and have no theoretical upper bound.

Several technical issues and measurement errors, occurring when the different fixed gain images are merged at NOAA, produce a lot of variability in the radiance-calibrated data: *i)* the low amplification data are based on considerably fewer orbits than the stable lights series (often covering only small parts of a year), *ii)* they are generated by blending different parts of the frequency spectrum which are deemed reliable, *iii)* higher light intensities are supported by fewer and fewer fixed-gain images[1], and *iv)* fires or stray lights are not fully removed from the auxiliary data. All this contributes to the high variance across different radiance-calibrated satellite-years.[2] Because of this instability, together with the fact that they are only available for seven out of 22 years, we only rely on the radiance-calibrated data to infer the shape of the distribution at the top. The relative ranks of pixels are consistently measured across the different satellites and less prone to be affected by measurement errors.

Table A-1 reports summary statistics for the 34 stable light satellite years and the seven radiance-calibrated years. Between 2.7% and 5.9% of all pixels in the stable lights images reach the top of the scale (i.e., 55 DN to 63 DN), more so in later years. As the radiance-calibrated lights do not suffer from top-coding, their mean, standard deviation, and Gini in lights are much higher. Rather than being capped at 63 DN, they reach maximum values from 2000 to 5000 DN. The fluctuations across satellites are reflected in the overall mean light intensity but are most apparent at the top. The maximum

---

[1]Consider the 2010 radiance-calibrated product. The maximum number of cloud-free images is 134, the suburbs of Paris are informed by about 50–60 cloud-free images, but the city core only by 10–20 images. This pattern repeats itself throughout all major cities.

[2]Measurement errors are also present in the stable lights data and affect their reliability in the time-series dimension but to a much lesser extent. The sensors of the satellites deteriorated over their lifetime and had to be replaced every couple of years, which implies that later recordings of any particular satellite tend to be the brightest (although this is not a hard rule). In panel regressions, economists usually resort to a combination of satellite and time fixed effects to partially address this issue.

light intensity doubles within three years and then decreases again by a similar amount (whereas the mean increases and decreases by about 27% over the same period).

Table A-2 confirms that these fluctuations are not driven by a few outliers. Instead of examining overall maxima, we now report various percentiles for the seven satellite-years of radiance-calibrated data and the means above these percentiles. For example, the top 2% begin at 147.01 DN in the 1996 data, at 214.59 DN in 2003, and again at 150.90 DN in 2010. The means above the various percentiles also vary over time. The differences are largest in absolute values at the very top but remain sizable throughout the distribution. We attribute this variation to the fact that fewer and fewer daily images inform the top of the distribution.

Table A-3 illustrates that not all differences between the stable lights and radiance-calibrated data can be attributed to top-coding. It regresses all pixels below 55 DN of the stable lights on the radiance-calibrated lights, where top-coding is supposed to not play a role. We find a regression coefficient around one-half rather than equivalence. This absence of a one-to-one correspondence is owed to the lack of onboard calibration, blooming (Abrahams et al., 2018), the presence of stray light (Hsu et al., 2015), and geo-location errors (Tuttle et al., 2013).

Table A-4 shows the maximum values attained by the seven satellite-years of radiance-calibrated data in 30 selected cities. Despite considerable variability over time, the relative ranking is in line with our expectations. The light intensity of the brightest pixel in New York City, for example, is about ten times greater than that of the brightest pixel in Nairobi. Note that the average maximum light intensity hardly exceeds 2000 DN, no matter if we compute it for London, New York, or Shanghai. This is why we restrict the maximum light intensities generated by our pixel-level correction to 2000 DN. Alternatively, our approach can be interpreted as fixing the radiance-calibrated scale at its 2010 or 2000 range, as these are the two years in which no city pixel exceeds 2000 DN.

Table A-5 reports the maximum light intensities recorded within 25 kilometers of the city center in 988 world cities with more than 500,000 inhabitants. Table A-6 adds the rank-correlations. The latter are much higher and typically around 0.90–0.95 for adjacent radiance-calibrated years, which supports our preference for pixel ranks over their actual values.

Table A-1 – Summary statistics of the stable lights and radiance-calibrated data

Year	Flight No.	Mean	Std. Dev.	Gini	% ≥ 55	Mean	Std. Dev.	Gini	Max
		Stable lights				*Radiance-calibrated*			
1992	F10	13.83	13.51	0.44	3.81				
1993	F10	11.96	12.81	0.46	3.12				
1994	F10	12.02	13.31	0.48	3.49				
	F12	14.65	13.93	0.44	4.20				
1995	F12	13.09	13.57	0.46	3.76				
1996	F12	12.69	13.36	0.46	3.51	19.42	55.63	0.65	2064
1997	F12	13.45	13.74	0.45	3.94				
	F14	10.98	12.87	0.49	3.16				
1998	F12	13.89	13.89	0.45	4.18				
	F14	10.94	12.78	0.49	3.05				
1999	F12	14.74	14.34	0.44	4.67	19.53	56.93	0.64	4698
	F14	10.15	12.31	0.49	2.78				
2000	F14	11.34	12.99	0.49	3.18	22.88	65.84	0.63	5552
	F15	13.25	13.34	0.44	3.70				
2001	F14	11.64	13.32	0.49	3.50				
	F15	12.93	13.26	0.45	3.54				
2002	F14	12.14	13.70	0.49	3.77				
	F15	13.18	13.44	0.45	3.72				
2003	F14	11.96	13.72	0.49	3.82	24.83	67.57	0.65	4186
	F15	10.28	12.45	0.50	2.70				
2004	F15	10.08	12.52	0.51	2.76	24.07	65.94	0.66	4357
	F16	11.82	13.04	0.46	3.40				
2005	F15	10.44	12.73	0.51	2.79				
	F16	10.44	12.54	0.49	2.85				
2006	F15	10.56	12.91	0.51	2.93	20.63	50.93	0.63	3333
	F16	12.26	13.37	0.47	3.48				
2007	F15	10.74	12.82	0.50	2.79				
	F16	13.05	13.79	0.46	4.03				
2008	F16	12.97	13.84	0.47	3.95				
2009	F16	13.50	14.12	0.47	4.17				
2010	F18	17.55	15.35	0.43	5.91	19.04	44.35	0.60	2110
2011	F18	14.78	14.68	0.46	4.94				
2012	F18	16.44	15.20	0.44	5.76				
2013	F18	16.23	15.20	0.44	5.78				

Notes: The table reports summary statistics using a 10% sample of the stable lights and radiance-calibrated data at the pixel level, where each pixel is 30 × 30 arc seconds. There are several years when two DMSP satellites were concurrently recording data for the stable lights series, so that there are 34 satellite-years between 1992 and 2013. The radiance-calibrated data are only available for the following periods: 16 Mar 96 – 12 Feb 97 (1996), 19 Jan 99 – 11 Dec 99 (1999), 03 Jan 00 – 29 Dec 00 (2000), 30 Dec 02 – 11 Nov 2003 (2003), 18 Jan 04 – 16 Dec 04 (2004), 28 Nov 05 – 24 Dec 06 (2006), and 11 Jan 10 – 9 Dec 10 (2010), although the actual coverage in terms of days often refers to a much smaller period.

Table A-2 – Summary statistics of the radiance-calibrated data (top shares)

Year	1996	1999	2000	2003	2004	2006	2010
				Panel a) Top 5%			
Percentile (x)	62.87	66.74	73.75	94.60	90.97	74.97	64.84
Mean above x	186.04	197.42	228.61	245.59	236.80	189.34	166.90
				Panel b) Top 4%			
Percentile (x)	76.30	84.79	95.62	119.27	114.26	94.03	81.98
Mean above x	215.29	228.01	264.84	280.40	270.51	215.70	190.42
				Panel c) Top 3%			
Percentile (x)	98.42	114.12	131.13	154.40	149.82	122.72	108.27
Mean above x	258.23	271.33	315.97	328.70	317.06	251.82	222.49
				Panel d) Top 2%			
Percentile (x)	147.01	166.33	198.77	214.59	207.97	168.84	150.90
Mean above x	327.60	338.23	393.22	402.32	387.54	305.83	269.84
				Panel e) Top 1%			
Percentile (x)	259.04	275.41	318.53	331.88	314.53	255.44	229.79
Mean above x	460.17	463.60	534.81	538.85	519.98	404.80	354.36
				Panel f) Top 0.1%			
Percentile (x)	729.41	716.94	815.16	822.00	805.43	605.13	511.62
Mean above x	979.91	960.86	1117.96	1110.62	1111.93	806.63	687.53
				Panel g) Top 0.01%			
Percentile (x)	1355.38	1279.48	1528.71	1491.25	1516.16	1085.71	936.22
Mean above x	1551.16	1652.31	1893.03	1828.03	1914.32	1316.93	1137.76

Notes: The table shows summary statistics of the radiance-calibrated data at the various percentiles. The input data are a 10% representative sample of all non-zero lights in the radiance-calibrated data above the defined threshold at the pixel level, where each pixel is 30 × 30 arc seconds.

Table A-3 – Regression of stable lights on radiance-calibrated data

Year	1996	1999	2000	2003	2004	2006	2010
Stable lights	0.5557	0.5502	0.4241	0.4357	0.3473	0.4874	0.7468
	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0004)
Constant	4.6422	5.5218	3.9172	3.6007	3.2069	2.9392	6.4094
	(0.0045)	(0.0054)	(0.0044)	(0.0043)	(0.0036)	(0.0037)	(0.0066)
R^2	0.7440	0.7013	0.7115	0.7709	0.7873	0.8011	0.6319

Notes: The table reports OLS estimates of a regression of all pixels smaller than 55 DN of the stable lights on their radiance-calibrated counterpart in all those years for which both data sources are available. Standard errors are in parentheses. The data are a 10% random sample of lights at the pixel level, where each pixel is 30 × 30 arc seconds.

Table A-4 – Maximum light intensities in 30 selected cities over time

City	1996	1999	2000	2003	2004	2006	2010	Average
Beijing	977.94	2265.07	3160.86	2911.25	2979.75	1575.00	1262.30	2161.74
Berlin	393.16	904.28	648.10	883.65	1045.85	474.82	357.46	672.48
Bogota	416.75	661.73	774.02	828.82	602.80	622.18	489.89	628.02
Brussels	882.44	1026.01	1410.89	840.66	935.03	920.00	465.11	925.73
Cairo	1785.16	1709.40	1768.79	2013.90	1876.90	1273.33	940.31	1623.97
Calgary	1669.85	1084.76	2077.92	1520.70	822.00	731.15	721.22	1232.51
Casablanca	729.69	919.44	769.98	1075.77	1214.73	708.33	620.97	862.70
Damascus	952.80	1302.40	921.07	862.37	1068.60	852.94	766.55	960.96
Dhaka	321.94	427.63	370.72	458.76	438.15	296.59	208.28	360.29
Dubai	1882.37	2137.78	2144.43	2068.70	2104.32	1457.14	1169.82	1852.08
Edinburgh	453.18	811.04	537.69	767.20	973.48	425.24	518.75	640.94
Foshan	537.46	715.98	1410.36	1625.73	1499.66	1142.86	1164.98	1156.72
Istanbul	379.97	1018.80	779.05	681.09	902.42	652.94	743.44	736.82
Jakarta	1100.27	683.82	664.56	788.43	1381.62	805.95	632.81	865.35
Johannesburg	517.62	672.89	712.18	688.42	729.53	528.57	448.43	613.95
London	984.36	2332.85	1940.67	1664.98	1356.30	1111.11	511.40	1414.52
Los Angeles	1214.68	1331.26	1805.43	1661.61	1561.80	1153.33	1199.81	1418.27
Manila	629.67	629.42	695.65	810.81	808.30	580.00	513.82	666.81
Moscow	976.78	1202.94	1496.44	1729.90	2383.80	1250.00	1285.51	1475.05
Mosul	136.55	139.61	161.64	194.56	244.10	85.71	148.04	158.60
Mumbai	527.60	543.47	622.09	753.50	730.67	550.00	414.87	591.74
Nairobi	211.83	180.27	188.66	191.45	173.54	174.02	164.13	183.41
New York	1664.97	3342.95	2145.71	2123.50	1575.50	1815.38	1366.45	2004.92
Paris	1177.72	1827.80	2444.32	1794.70	1430.28	1425.00	874.55	1567.77
Rio de Janeiro	748.92	926.51	917.27	708.83	699.31	484.08	461.57	706.64
Seoul	1327.49	1847.79	1999.46	2049.13	2285.16	1735.71	924.85	1738.51
Shanghai	1123.89	1965.24	1906.01	3982.13	2931.80	2307.14	1926.59	2306.12
Sydney	1070.85	1138.12	954.50	1223.82	1262.36	840.00	513.78	1000.49
Tel Aviv	997.83	1284.19	1679.72	1446.72	1397.40	1188.24	1099.83	1299.13
Tokyo	1515.98	1456.54	1790.01	1963.82	1775.52	1322.22	1842.57	1666.67

Notes: The table report the maximum light intensity in DN recorded within 25 km radius of the city center in a selection of cities. The input data are the radiance-calibrated lights. City locations are obtained from the Natural Earth point data of major populated places.

Table A-5 – Correlation matrix of maximum city lights

Years	1996	1999	2000	2003	2004	2006	2010
1996	1.0000						
1999	0.8612	1.0000					
2000	0.8551	0.9134	1.0000				
2003	0.8065	0.8733	0.8990	1.0000			
2004	0.7713	0.8587	0.8601	0.9176	1.0000		
2006	0.8106	0.8741	0.8953	0.9379	0.9305	1.0000	
2010	0.7406	0.7806	0.7839	0.8501	0.8513	0.8939	1.0000

Notes: The table reports correlations between the maximum light intensities recorded within 25 km radius of the city center of 988 world cities with more than 500,000 inhabitants.

Table A-6 – Rank correlation matrix of maximum city lights

Years	1996	1999	2000	2003	2004	2006	2010
1996	1.0000						
1999	0.9161	1.0000					
2000	0.9126	0.9557	1.0000				
2003	0.8448	0.9129	0.9328	1.0000			
2004	0.8453	0.9048	0.9162	0.9542	1.0000		
2006	0.8621	0.9066	0.9253	0.9548	0.9491	1.0000	
2010	0.8096	0.8489	0.8645	0.8970	0.8968	0.9270	1.0000

Notes: The table reports rank correlations between the maximum light intensities recorded within 25 km radius of the city center of 988 world cities with more than 500,000 inhabitants.

B The top-coding threshold

The influence of top-coding in the DMSP-OLS satellite data has been underestimated in part because much of the literature assumes it only affects pixels with the highest recorded value. However, even though the scale of stable lights goes up to 63, we have good reason to assume that many pixels with DNs of 62, 61, down to the mid-50s, are subject to top-coding and should be brighter than they are recorded in the data.

The rationale behind this conjecture is straightforward. The stable lights data were already averaged at least twice during the data construction. First, the DMSP satellites averaged several higher resolution pixels on-board to reduce the amount of information that needs to be transmitted down to Earth. The OLS system recorded images at a nominal resolution of 0.56 km, which was averaged on-board into 5×5 blocks to create a 2.77 km (smooth) resolution and then reprojected onto a 30 arc-second grid.[3] Second, the data providers at NOAA processed the daily images into a single annual composite. As a result, many pixels suffering from top-coding in at least one of the underlying fine resolution data points or smooth resolution daily images would have ended up with an average value of less than 63. Hsu et al. (2015) suggest that this subtle type of top-coding may even start at a DN as low as 35. Since "the OLS does onboard averaging to produce its global coverage data, saturation does not happen immediately when radiance reaches the maximum level. On the contrary, as the actual radiance grows, the observed DN value fails to follow the radiance growth linearly, causing a gradual transition into a plateau of full saturation" (Hsu et al., 2015, p. 1872).

We explore the location of the top-coding threshold with a statistical approach. If only the stable lights at 63 DN were subject to top-coding, we would expect the histogram in panel (a) of Figure B-1 to show a decreasing shape ending in a spike only at 63 DN. Instead, we observe an increase in the number of pixels from 55 onwards (e.g. a bathtub shape), signaling that these values are top-coded as well. Further evidence along these lines is provided by panel (b) of Figure B-1. It shows a histogram of the light intensity of the stable lights DNs associated with high radiance-calibrated values (above 160 DN). There are a large number of pixels with DNs down to the mid-50s which correspond to very high radiance-calibrated values, but the density falls rapidly below the mid-50s. Other years show very similar patterns.

Table B-1 list the percentile values of the radiance-calibrated lights corresponding to stable lights at 55 DN, 56 DN, and so on. The stable lights at 63 DN have the highest radiance-calibrated values (50% of them are higher than 390 DN). But there is also a significant share of 55 DN lights corresponding to high radiance-calibrated values, for instance, 25% are recorded with 140 DN or brighter.

---

[3]See https://directory.eoportal.org/web/eoportal/satellite-missions/d/dmsp-block-5d or Abrahams et al. (2018) for a detailed description of the sensors and on-board processing.

Figure B-1 – Histograms of stable lights in 1999

(a) If stable DN > 9   **(b)** If radiance-calibrated DN > 160
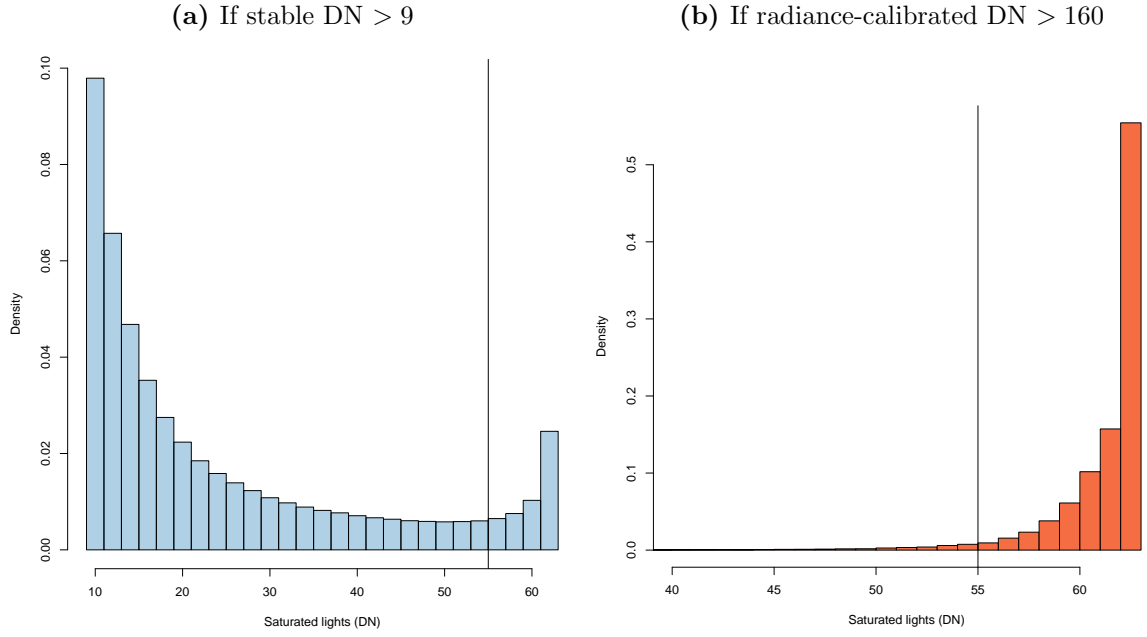
Notes: Illustration of the location of the top-coding threshold in the stable lights. Panel a) shows a histogram of the F12 satellite in 1999 for all pixels with a DN greater 9. Panel b) shows a histogram of the same satellite only for pixels where the radiance-calibrated light intensity is greater 160 DN. The input data are a 10% representative sample of all non-zero lights in the stable lights and radiance-calibrated data at the pixel level (see Elvidge et al., 2009, Hsu et al., 2015).

Table B-1 – Percentiles of radiance-calibrated values at given stable lights values in 2000

| Stable lights | Radiance-calibrated percentiles | | | | | |
DN	5%	25%	50%	75%	95%	99%
55	53.20	74.94	99.41	140.85	232.90	328.86
56	56.15	79.99	108.20	153.92	250.93	344.05
57	60.14	84.99	115.11	164.63	262.18	357.60
58	64.13	92.81	125.35	179.57	277.59	392.33
59	70.32	101.97	141.92	203.17	306.77	423.28
60	79.16	116.64	163.92	231.91	344.57	497.25
61	89.33	137.89	196.68	268.21	410.91	625.30
62	109.03	176.36	246.66	331.46	524.18	762.63
63	160.91	276.92	390.08	560.28	952.14	1494.85

Notes: The table reports values from the cumulative distribution function of the radiance-calibrated lights which are associated with a given stable lights value (from 55 to 63). For instance, 25% of the radiance-calibrated values associated with a stable lights value of 61 DN, are below 137.89. The data are a representative 10% sample for the year 2000.

C   Additional results for the model

C.1   Proofs

In this section, we provide additional proofs of the model presented in the main text.

The CDF of the number of rings: Note that $r = \pi^{-1/2}x^{1/(2\phi)}$ implies $x = \pi^\phi r^{2\phi}$ and $\mathrm{d}x = 2\phi\pi^\phi r^{2\phi-1}\mathrm{d}r$. Substituting these definitions into eq. (1) and integrating yields the CDF of the number of rings per city as presented in eq. (2) of the main text

$$
\begin{aligned}
F(r) &= 2\phi x_c \pi^{-\phi} \int_{\tilde{r}}^{r} r^{-2\phi-1}\mathrm{d}r = 2\phi x_c \pi^{-\phi}\left[-\frac{1}{2\phi}r^{-2\phi}\right]_{\tilde{r}}^{r} \\
&= \begin{cases} 0 \text{ for } r < \tilde{r} = \pi^{-1/2}x_c^{1/(2\phi)} \\ 1 - y_c\pi^{-\phi}r^{-2\phi} \text{ for } r >= \tilde{r} = \pi^{-1/2}x_c^{1/(2\phi)}. \end{cases}
\end{aligned}
\tag{C-1}
$$

The density of pixels: Start with the distribution of the number of pixels. At distances $d < \tilde{d}$, the amount of pixels increases linearly in d as rings farther away from the center contain more pixels: $\frac{\mathrm{d}}{\mathrm{d}d}\pi d^2 = 2\pi d$. Beyond $\tilde{d}$, the effect within each city has to be multiplied by the survival function $1 - F(r)$ from eq. (2), as there are fewer and fewer cities of such size. Denoting the number of cities as M, the absolute amount of pixels N as a function of d is

$$
P(d) = \begin{cases} 2\pi d M \text{ for } d < \tilde{d} = \pi^{-1/2}x_c^{1/(2\phi)} \\ 2\pi^{1-\phi}M x_c d^{1-2\phi} \text{ for } d \geq \tilde{d} = \pi^{-1/2}x_c^{1/(2\phi)}. \end{cases}
\tag{C-2}
$$

The total number of pixels, N, can be obtained by integration

$$
\begin{aligned}
N &= \int_0^{\tilde{d}} 2\pi d M \mathrm{d}d + \int_{\tilde{d}}^{\infty} 2\pi^{1-\phi}M x_c d^{1-2\phi}\mathrm{d}d = 2\pi M\left[\frac{1}{2}d^2\right]_0^{\tilde{d}} + 2\pi^{1-\phi}M x_c\left[\frac{1}{2-2\phi}d^{2-2\phi}\right]_{\tilde{d}}^{\infty} \\
&= \pi M\frac{y_c^{1/\phi}}{\pi} + \frac{\pi^{1-\phi}M y_c}{\phi-1}\left(\frac{y_c^{1/\phi}}{\pi}\right)^{1-\phi} = M x_c^{1/\phi} + \frac{1}{\phi-1}M x_c^{1/\phi} = \frac{\phi}{\phi-1}M x_c^{1/\phi}.
\end{aligned}
\tag{C-3}
$$

Dividing eq. (C-2) by N yields the density, $f(d)$, shown in eq. (4):

$$
f(d) = \begin{cases} 2\pi\frac{\phi-1}{\phi}x_c^{-1/\phi}d \text{ for } d < \tilde{d} \\ 2\pi^{1-\phi}\frac{\phi-1}{\phi}x_c^{1-1/\phi}d^{1-2\phi} \text{ for } d \geq \tilde{d} \end{cases}
\tag{C-4}
$$

with $\tilde{d} = \pi^{-1/2}x_c^{1/(2\phi)}$.

The density is illustrated in panel (a) of Figure 4 in the main text.

C.2 The maxima distribution of city lights

When assuming that lights within cities follow the negative exponential distribution (assumption 4b), our expression for the overall top light distribution depends on the distribution of light maxima across cities. We discuss three cases for the distribution of light maxima, which is informed by empirical evidence on the distribution of the maxima of 988 world cities (above 500,000 inhabitants), measured by the radiance-calibrated lights in 2010. We will here provide further empirical tests to underpin case 3 that this distribution is Pareto in the tail.

Starting with the histogram in Figure 5, we see a downward-sloping shape from around luminosity values of 500. We will therefore test for a Pareto tail using five different thresholds (500, 600, 700, 800, 900).

We follow Cirillo (2013) in drawing a discriminant moment ratio plot to provide evidence of the Pareto property. Figure C-1 plots the coordinate pair of the coefficient of variation (i.e., standard deviation divided by the mean) on the x-axis and skewness on the y-axis. As each parametric distribution has its particular curve of feasible coordinates, the plane can be divided into a Pareto area (comprising Pareto type I and II), a log-normal area and a gray area possibly belonging to both. Here we see that no matter which threshold we choose, the distribution of maximum light intensities across cities falls in the Pareto area.

Next, we estimate the Pareto coefficient of the maxima distribution. We run log-rank regressions with both the OLS and Hill estimator for the different thresholds, following the literature (see for instance Gabaix and Ibragimov, 2011). As Table C-1 shows, the Pareto coefficient is rather high, ranging from 2.6 to 3.5 depending on the threshold and estimator. While being Pareto, light maxima across cities are therefore rather equally distributed, compared to, for example, the overall light distribution. Also, we note that there is a slight increase in the Pareto α as the threshold increases and there are fewer observations left. But overall, the magnitude of the coefficients is rather stable.

Figure C-1 – Discriminant moment ratio plot for distribution of maxima
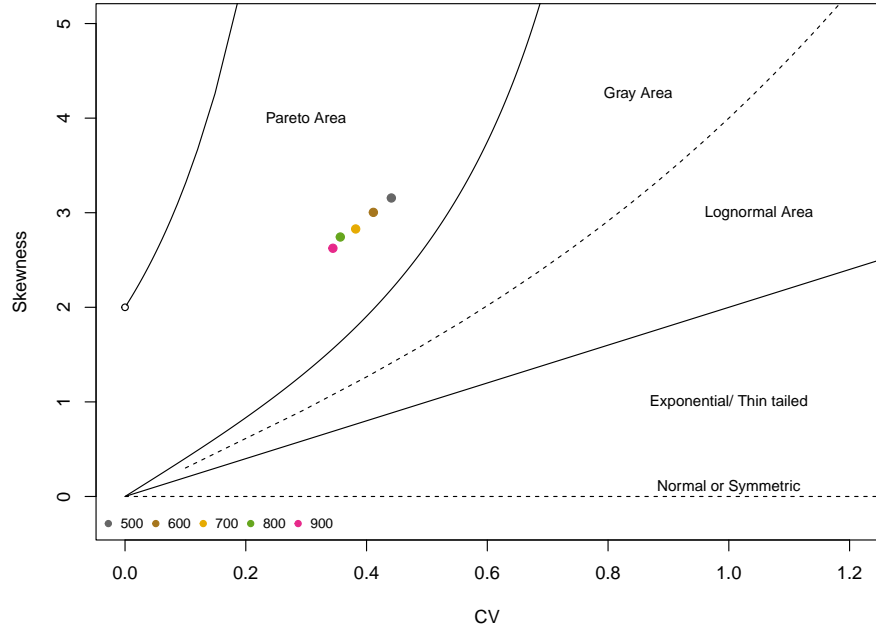
Notes: The panels show discriminant moment ratio plots (Cirillo, 2013) using various thresholds. The data are the distribution of maximum light intensities in 988 world cities above 500,000 inhabitants, measured by the radiance-calibrated satellite in 2010.

Table C-1 – Rank regressions for the maxima distribution

Threshold	500	600	700	800	900
Panel a) OLS Estimator					
Pareto $\hat{\alpha}$	2.6228	2.9077	2.9177	2.9955	3.3901
(S.E.)	(0.1578)	(0.2132)	(0.2675)	(0.3308)	(0.4272)
Observations	278	187	120	83	64
Panel b) Hill Estimator					
Pareto $\hat{\alpha}$	2.7787	3.0173	3.0917	3.2237	3.5042
(S.E.)	(0.0165)	(0.0183)	(0.0273)	(0.0347)	(0.0294)
Observations	278	187	120	83	64

Notes: The table reports the results of rank regressions with $\log\left(\text{rank}(y_i) - 1/2\right) - \log N$ as the dependent variable. Asymptotic standard errors computed as $(2/N)^{1/2}\hat{\alpha}$ are reported in parentheses (see Gabaix and Ibragimov, 2011).

D Extreme value theory

As an alternative to our stylized urban economics model, we can also motivate a Pareto distribution in top lights purely on statistical grounds using extreme value theory (EVT). EVT deals with the probability distributions of sparse observations such as threshold exceedances. A key result of this theory is that these quantities observe a Generalized Pareto distribution (Coles, 2001).

More precisely, let $X_1, X_2, \ldots$ be a sequence of independent random variables—such as light—with common but unknown distribution function F, and let $M_n = \max\{X_1, \ldots, X_n\}$. If F satisfies the *extremal types theorem* (Coles, 2001), so that for large n, $\mathbb{P}[M_n > z] \approx G(z)$ with $G(z)$ as the Generalized Extreme Value distribution, then, for a high enough threshold u, the distribution of the threshold exceedance $\mathbb{P}[(X - u) > y | X > u]$ is approximately

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}}, \tag{D-1}$$

no matter which regular distribution X was drawn from.

This means that we will observe a Generalized Pareto distribution with parameters ξ and σ for all lights values above a specified threshold. With $\xi = 0$, this reduces to the exponential distribution and with $\xi > 0$ the distribution is Pareto. There is strong evidence that the latter case holds for the lights data.

Figure D-1 plots the Generalized Pareto distribution against the empirical distribution function of the radiance-calibrated data from 2010. It visualizes the close fit.

Table D-1 shows the results of fitting the Generalized Pareto distribution to various top shares of the light distribution of the seven satellite-years of radiance-calibrated data. The fit is very good and the estimated ξ parameters are always significantly positive. This clearly points towards a Pareto distribution.

Figure D-1 – Generalized Pareto CDF versus EDF, radiance-calibrated data in 2010
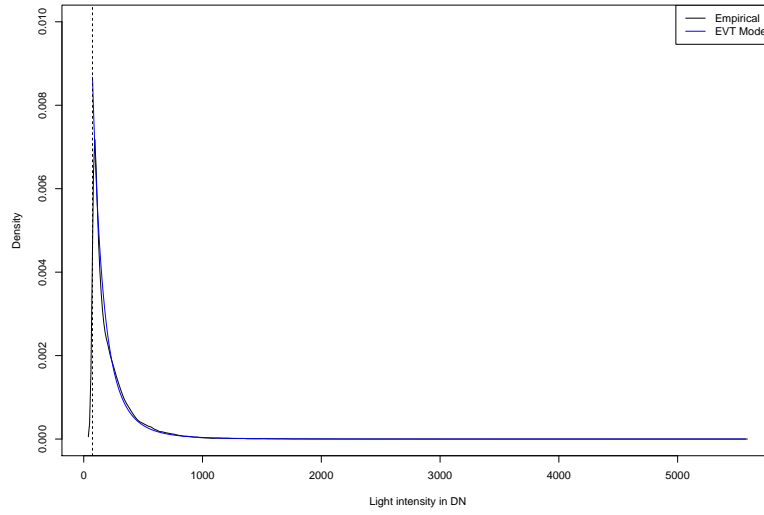
Notes: Illustration of Generalized Pareto CDF fitted to the data and the empirical distribution function (EDF). The EDF and Generalized Pareto CDF are fitted to the top 4% of stable lights in 2010. The input data are a 10% representative sample of all non-zero lights of the radiance-calibrated data at the pixel level, where each pixel is 30×30 arc seconds.

Table D-1 – Fitted Generalized Pareto distributions, varying thresholds

Year	1996	1999	2000	2003	2004	2006	2010	Average
				Panel a) Top 5%				
$\ln \sigma$	4.2422	4.5197	4.7304	4.7560	4.7136	4.5239	4.4477	4.5619
	(0.0062)	(0.0051)	(0.0053)	(0.0052)	(0.0050)	(0.0051)	(0.0049)	[0.1860]
ξ	0.4917	0.3136	0.2792	0.2354	0.2398	0.1972	0.1652	0.2746
	(0.0056)	(0.0043)	(0.0044)	(0.0042)	(0.0040)	(0.0040)	(0.0038)	[0.1075]
Threshold	63	67	74	95	91	75	65	–
Observations	96,685	116,858	106,914	100,094	106,899	99,486	107,745	–
				Panel b) Top 4%				
$\ln \sigma$	4.5136	4.6989	4.9099	4.8569	4.8241	4.6278	4.5574	4.7127
	(0.0066)	(0.0055)	(0.0055)	(0.0057)	(0.0054)	(0.0055)	(0.0052)	[0.1545]
ξ	0.3720	0.2401	0.2020	0.2051	0.2051	0.1605	0.1212	0.2152
	(0.0057)	(0.0044)	(0.0044)	(0.0045)	(0.0042)	(0.0043)	(0.0039)	[0.0790]
Threshold	76	85	96	119	114	94	82	–
Observations	77,348	93,484	85,481	80,075	85,489	79,589	86,195	–
				Panel c) Top 3%				
$\ln \sigma$	4.8260	4.8674	5.0702	4.9841	4.9127	4.7153	4.6387	4.8592
	(0.0069)	(0.0060)	(0.0060)	(0.0063)	(0.0061)	(0.0062)	(0.0058)	[0.1491]
ξ	0.2266	0.1753	0.1387	0.1629	0.1873	0.1356	0.0944	0.1601
	(0.0056)	(0.0047)	(0.0045)	(0.0049)	(0.0047)	(0.0047)	(0.0042)	[0.0424]
Threshold	98	114	131	154	150	123	108	–
Observations	58,010	70,112	64,110	60,057	64,133	59,691	64,646	–
				Panel d) Top 2%				
$\ln \sigma$	5.0520	5.0042	5.1151	5.0807	4.9771	4.7847	4.6836	4.9568
	(0.0079)	(0.0070)	(0.0073)	(0.0076)	(0.0075)	(0.0075)	(0.0070)	[0.1614]
ξ	0.1355	0.1332	0.1438	0.1432	0.1933	0.1266	0.0903	0.1380
	(0.0060)	(0.0053)	(0.0055)	(0.0057)	(0.0058)	(0.0056)	(0.0050)	[0.0304]
Threshold	147	166	199	215	208	169	151	–
Observations	38,673	46,742	42,740	40,039	42,755	39,795	43,097	–
				Panel e) Top 1%				
$\ln \sigma$	5.2035	5.1025	5.2287	5.1729	5.1013	4.8650	4.7009	5.0535
	(0.0109)	(0.0099)	(0.0103)	(0.0108)	(0.0109)	(0.0108)	(0.0100)	[0.1936]
ξ	0.0961	0.1262	0.1374	0.1483	0.2037	0.1324	0.1163	0.1372
	(0.0082)	(0.0074)	(0.0078)	(0.0082)	(0.0086)	(0.0082)	(0.0074)	[0.0337]
Threshold	259	275	319	332	315	255	230	–
Observations	19,337	23,371	21,370	20,019	21,378	19,897	21,548	–

Notes: The table reports parameter estimates from fitting the Generalized Pareto distribution shown in eq. (D-1). The input data are a 10% representative sample of all non-zero lights in the radiance-calibrated data above the defined threshold at the pixel level, where each pixel is 30×30 arc seconds. The last column reports the point average of the seven satellites and its standard deviation in brackets.

E   Additional results using the radiance-calibrated data

This section complements the analysis in the paper by providing additional robustness checks of our Pareto hypothesis using the seven satellite-years of radiance-calibrated data.

Visual Inspection: Panel (a) of Figure E-1 shows Zipf plots for the top 2% of lights for each of the seven satellite-years of radiance-calibrated data. A Zipf plot is a visualization of the Pareto survival function in logs. A linear Zipf plot is usually considered evidence in favor of the Pareto distribution, but its practical relevance is being contested (Cirillo, 2013). Our plots for the lights data are qualitatively similar to those of the top incomes literature, in that they display linear sections together with some initial curvature and outliers at the end.[4] It is well-known that Zipf plots often deviate from linearity at the very top since fewer and fewer values are observed at the extremes. Sometimes this is addressed by removing the very top. We use logarithmic bins so that the size of the bins increases by a multiplicative factor (Newman, 2005). The sensitivity of Zipf plots to outliers is compounded by instability and measurement errors afflicting the radiance-calibrated satellites. While we conclude that the Zipf plot using the radiance-calibrated data is ambiguous, we obtain a near-linear Zipf plot using the superior VIIRS data (see the next section).

Panel (b) of Figure E-1 provides another graphical test for the Pareto distribution based on 'Van der Wijk's Law'. The Pareto distribution is unique in that the average above some level y is proportional to y at all points in the tail, with a factor of proportionality equal to $\frac{\alpha}{\alpha-1} > 1$. The graph plots, for each DN on the x-axis, the average luminosity of all pixels brighter than this value on the y-axis. As expected, we observe a linear relationship with a slope above unity.

Tests against the log-normal distribution: As a robustness check, we pit the Pareto distribution against other plausible candidates. We pay particular attention to the log-normal distribution, since it is commonly used to describe the complete distribution of incomes or city sizes. Table E-1 shows the results from separate regressions of the empirical distribution function on the Pareto CDF and the log-normal CDF based on the top 4% of the data. The estimated coefficient for the Pareto CDF is closer to unity and the R^2 is substantially larger than in the log-normal counterpart (0.98 vs. 0.83). Figure E-2 visualizes this difference in fit for the year 2010. The log-normal CDF fits the data poorly, while the Pareto CDF is always closer to the empirical distribution.

---

[4]Working with any top share, from the top 5% to the top 1% gives qualitatively similar results, even if the case for a Pareto distribution tends to be stronger the higher we set the threshold. This is in line with the empirical literature on Pareto applications in other fields.

Unrestricted rank regressions: Recall that for Pareto-distributed observations y_i, $i = 1, ... N$, we have $\text{rank}(y_i) \approx N y_c^\alpha y_i^{-\alpha}$, or, in logarithms $\log \text{rank}(y_i) - \log N \approx \alpha \log y_c - \alpha \log y_i$. Hence, in the regression

$$\log\left(\text{rank}(y_i) - \frac{1}{2}\right) - \log N = \alpha_1 \log y_c + \alpha_2 \log y_i + \epsilon \qquad \text{(E-1)}$$

only the Pareto distribution satisfies the null hypothesis that $-\alpha_1 = \alpha_2$ with $\alpha_2 < 0$. As before, we follow Gabaix and Ibragimov (2011) and subtract one half from the rank to improve the OLS estimation of the tail exponent in the rank regression.

Table E-2 reports the OLS rank regression results of eq. (E-1) for all seven satellites at various different thresholds, i.e. the top 5% to top 1%. The two coefficients are usually very close and the R^2s are high (0.96–0.99).[5]

The Hill estimator: If the null hypothesis $-\alpha_1 = \alpha_2 = \alpha$ is enforced in eq. (E-1), one can directly obtain the parameter estimate for the Pareto α. In the main text we estimate this parameter using OLS rank regressions. As a robustness check, we now use the Hill estimator (Hill, 1975), $\hat{\alpha}_{Hill} = (N - 1) \left(\sum_{i=1}^{N-1} \log y_i - \log y_c \right)^{-1}$, for the restricted rank regression

$$\log \text{rank}(y_i) - \log N \approx \alpha \log y_c - \alpha \log y_i. \qquad \text{(E-2)}$$

Under the assumption of a Pareto distribution, the Hill estimator equals the efficient maximum likelihood estimator and is known for its superior properties for fitting the tail of the Pareto distribution (Eeckhout, 2009). The standard errors are given by $\hat{\alpha}_{Hill}/\sqrt{N-3}$ (see Gabaix, 2009).

Table E-3 report the results for all seven satellites at various different thresholds, i.e. the top 5% to top 1%. The Pareto parameters obtained using the Hill estimator are very similar to the OLS estimates in the main text. For the top 3-4%, the values are between 1.3 and 1.6 for the seven satellites, very close to the average OLS parameter estimate of 1.5. For higher thresholds, we observe also the same increase in the parameter estimate that we observe in the OLS results.

---

[5]Note that formal statistical tests, e.g. tests of coefficient equality or Kolmogorov-Smirnoff tests, do not make much sense in huge samples such as ours. Gabaix and Ioannides (2004, p. 2350) capture this nicely: "with an infinitely large dataset one can reject any non-tautological theory." The extremely small standard errors lead to overrejections of the null hypothesis unless the empirical value equals exactly the theoretical value.

Figure E-1 – Zipf plot and Van der Wijk's plot

(a) Zipf
(b) Van der Wijk
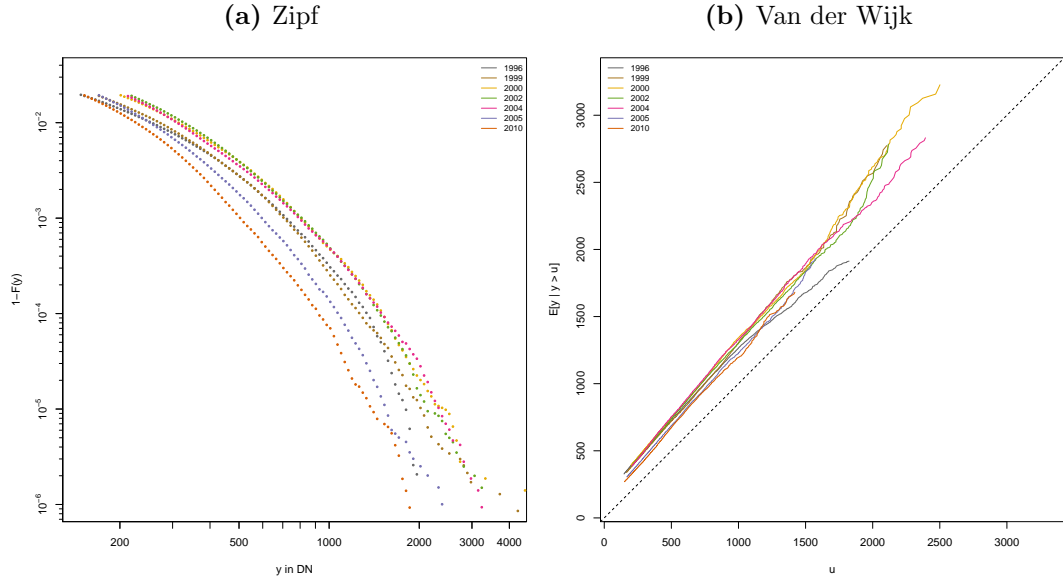
Notes: Popular graphical tests for an approximate Pareto distribution in top lights. Panel (a) shows the Zipf plot for the top 2% of all pixels. The figure uses logarithmic binning to reduce noise and sampling errors in the right tail of the distribution (see Newman, 2005). There are about 100 bins in the tail, where the exact number depends on the range of the input data. Panel (b) demonstrates Van der Wijk's law, which states that the average light above some value u is proportional to u, this is $E[y|y > u] \propto u$. Here, too, the data is the top 2% of all pixels. The input data are a 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level.

Figure E-2 – Pareto and log-normal CDF versus EDF, radiance-calibrated lights in 2010
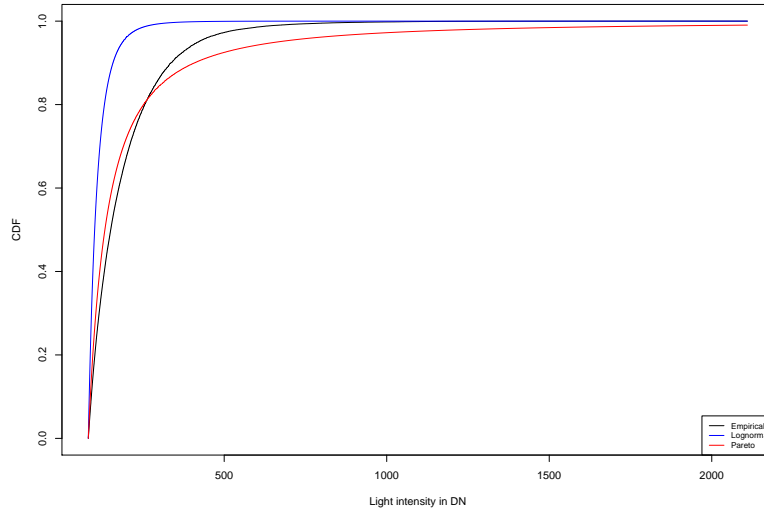
Notes: Illustration of the difference between the Pareto and log-normal CDFs fitted to the data and the empirical distribution function (EDF). Note that the log-normal distribution was fitted to the whole distribution rather than the tail because of its unimodal shape, while the Pareto distribution is estimated only on the tail. For comparison, we adjust the CDFs so that they all start at the top 4% of radiance-calibrated lights in 2010. The input data are a 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level, where each pixel is 30×30 arc seconds.

xviii

Table E-1 – Regression of the EDF on theoretical CDFs, top 4%

Year	1996	1999	2000	2003	2004	2006	2010	Average
			Panel a) Pareto CDF on RHS					
Slope	1.0108	1.0551	1.0616	1.0575	1.0746	1.0722	1.0796	1.0588
	(0.0003)	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0005)	[0.0231]
Constant	-0.0320	-0.0668	-0.0746	-0.0666	-0.0787	-0.0784	-0.0858	-0.0690
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	[0.0177]
R^2	0.9914	0.9866	0.9802	0.9884	0.9869	0.9860	0.9831	–
			Panel b) Log-normal CDF on RHS					
Slope	0.9004	0.9265	0.9181	0.9387	0.9520	0.9472	0.9488	0.9331
	(0.0014)	(0.0016)	(0.0018)	(0.0014)	(0.0014)	(0.0014)	(0.0015)	[0.0190]
Constant	-0.1653	-0.2186	-0.2238	-0.1954	-0.2088	-0.2031	-0.2179	-0.2047
	(0.0011)	(0.0013)	(0.0015)	(0.0011)	(0.0011)	(0.0011)	(0.0012)	[0.0200]
R^2	0.8496	0.7913	0.7626	0.8508	0.8467	0.8480	0.8268	–

Notes: The table reports results of a regression of the empirical distribution function (EDF) on the Pareto or log-normal CDF, using the top 4% of the data. The data are a 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level, where each pixel is 30×30 arc seconds. The last column reports the point average of the seven satellites and its standard deviation in brackets.

Table E-2 – Unrestricted rank regressions

Year	1996	1999	2000	2003	2004	2006	2010	Average
				Panel a) Top 5%				
y_i	-1.4334	-1.4996	-1.4630	-1.6903	-1.6933	-1.7388	-1.7170	-1.6050
	(0.0012)	(0.0012)	(0.0013)	(0.0013)	(0.0013)	(0.0014)	(0.0015)	[0.1330]
y_c	1.4736	1.5632	1.5318	1.7539	1.7582	1.8087	1.7936	1.6690
	(0.0014)	(0.0015)	(0.0016)	(0.0016)	(0.0015)	(0.0017)	(0.0018)	[0.1405]
R^2	0.9694	0.9651	0.9600	0.9701	0.9721	0.9677	0.9620	0.9666
Observations	96,685	116,858	106,914	100,095	106,899	99,487	107,745	104,955
				Panel b) Top 4%				
y_i	-1.5130	-1.6328	-1.6165	-1.8403	-1.8513	-1.9101	-1.9056	-1.7528
	(0.0015)	(0.0014)	(0.0016)	(0.0016)	(0.0014)	(0.0017)	(0.0018)	[0.1612]
y_c	1.5618	1.6974	1.6870	1.8978	1.9132	1.9767	1.9804	1.8163
	(0.0017)	(0.0017)	(0.0019)	(0.0018)	(0.0016)	(0.0019)	(0.0020)	[0.1655]
R^2	0.9662	0.9661	0.9623	0.9725	0.9759	0.9711	0.9658	0.9685
Observations	77,348	93,484	85,482	80,075	85,489	79,590	86,196	83,952
				Panel c) Top 3%				
y_i	-1.6609	-1.8385	-1.8624	-2.0491	-2.0633	-2.1470	-2.1746	-1.9708
	(0.0019)	(0.0018)	(0.0019)	(0.0019)	(0.0016)	(0.0020)	(0.0021)	[0.1882]
y_c	1.7225	1.9017	1.9340	2.1044	2.1174	2.2068	2.2429	2.0328
	(0.0022)	(0.0020)	(0.0022)	(0.0021)	(0.0018)	(0.0022)	(0.0024)	[0.1872]
R^2	0.9646	0.9695	0.9691	0.9761	0.9811	0.9762	0.9721	0.9727
Observations	58,011	70,115	64,111	60,058	64,134	59,692	64,647	62,967
				Panel d) Top 2%				
y_i	-1.9711	-2.1628	-2.2315	-2.3687	-2.3478	-2.4809	-2.5663	-2.3042
	(0.0025)	(0.0023)	(0.0022)	(0.0023)	(0.0018)	(0.0024)	(0.0025)	[0.2009]
y_c	2.0329	2.2180	2.2831	2.4156	2.3880	2.5295	2.6215	2.3555
	(0.0029)	(0.0025)	(0.0025)	(0.0025)	(0.0020)	(0.0026)	(0.0027)	[0.1974]
R^2	0.9698	0.9757	0.9798	0.9825	0.9871	0.9826	0.9807	0.9797
Observations	38,673	46,742	42,740	40,039	42,756	39,794	43,097	41,977
				Panel e) Top 1%				
y_i	-2.5471	-2.7216	-2.7241	-2.8508	-2.7006	-2.9769	-3.1652	-2.8123
	(0.0039)	(0.0031)	(0.0031)	(0.0030)	(0.0027)	(0.0032)	(0.0029)	[0.2049]
y_c	2.5922	2.7593	2.7596	2.8823	2.7258	3.0097	3.2005	2.8471
	(0.0043)	(0.0034)	(0.0034)	(0.0033)	(0.0029)	(0.0035)	(0.0031)	[0.2031]
R^2	0.9781	0.9849	0.9864	0.9889	0.9895	0.9886	0.9911	0.9868
Observations	19,337	23,373	21,372	20,020	21,377	19,898	21,551	20,990

Notes: The table reports OLS results obtained from the unrestricted rank regressions eq. (E-1) at various relative thresholds. The input data are a 10% representative sample of all non-zero lights in the radiance-calibrated data above the defined threshold at the pixel level, where each pixel is 30×30 arc seconds. Standard errors are in parentheses. The last column reports the point average of the seven satellites and its standard deviation in brackets.

Table E-3 – Parameter estimates from rank regressions (Hill estimator)

Year	1996	1999	2000	2003	2004	2006	2010	Average
				Panel a) Top 5%				
Pareto $\hat{\alpha}$	1.2286	1.1833	1.1289	1.3112	1.3100	1.3356	1.3012	1.2570
	(0.0040)	(0.0035)	(0.0035)	(0.0041)	(0.0040)	(0.0042)	(0.0040)	[0.0780]
Observations	96,685	116,858	106,914	100,095	106,899	99,487	107,745	–
				Panel b) Top 4%				
Pareto $\hat{\alpha}$	1.2487	1.2689	1.2233	1.4431	1.4315	1.4666	1.4333	1.3593
	(0.0045)	(0.0042)	(0.0042)	(0.0051)	(0.0049)	(0.0052)	(0.0049)	[0.1065]
Observations	77,348	93,484	85,482	80,075	85,489	79,590	86,196	–
				Panel c) Top 3%				
Pareto $\hat{\alpha}$	1.2948	1.4152	1.3805	1.6023	1.6234	1.6672	1.6478	1.5188
	(0.0054)	(0.0053)	(0.0055)	(0.0065)	(0.0064)	(0.0068)	(0.0065)	[0.1509]
Observations	58,011	70,115	64,111	60,058	64,134	59,692	64,647	–
				Panel d) Top 2%				
Pareto $\hat{\alpha}$	1.5068	1.6869	1.7536	1.8920	1.9325	1.9860	2.0095	1.8239
	(0.0077)	(0.0078)	(0.0085)	(0.0095)	(0.0093)	(0.0100)	(0.0097)	[0.1832]
Observations	38,673	46,742	42,740	40,039	42,756	39,794	43,097	–
				Panel e) Top 1%				
Pareto $\hat{\alpha}$	2.0363	2.2458	2.2613	2.4101	2.3582	2.5190	2.6558	2.3552
	(0.0146)	(0.0147)	(0.0155)	(0.0170)	(0.0161)	(0.0179)	(0.0181)	[0.2011]
Observations	19,337	23,373	21,372	20,020	21,377	19,898	21,551	–

Notes: The table reports the results of the restricted rank regression eq. (E-2) using the Hill estimator. The data are a 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level, where each pixel is 30×30 arc seconds. The last column reports the point average of the seven satellites and its standard deviation in brackets.

F   Additional results using the VIIRS data

Since October 2011, the first satellite of the Suomi National Polar Partnership Visible Infrared Imagining Radiometer Suite (NPP-VIIRS) has been in orbit. The VIIRS day-night-band (DNB) on-board sensors have a much higher native resolution of 15 arc seconds, are radiometrically calibrated, do not suffer from top-coding, and record a physical quantity (radiance). This section complements the analysis in the paper by providing additional robustness checks of our Pareto hypothesis using this new data.

Although the new system is undoubtedly superior in many respects, comparability with the previous series is limited for at least two reasons: *i)* the first annual VIIRS composite made available by NOAA refers to the year 2015, so that there is no temporal overlap with the 1992-2013 DMSP-OLS series, *ii)* the VIIRS satellites have an overpass time around midnight, in contrast to the evening hours of the DMSP-OLS satellites, so that it is not entirely clear what kind of production and consumption activity they capture (Elvidge et al., 2014, Nordhaus and Chen, 2015). While we do not rely on the VIIRS data for our replacement procedure, we use the first VIIRS cross-section from 2015 as another robustness check for whether the Pareto distribution holds. The VIIRS data are particularly insightful in this respect because of their superior quality.

To compare the higher resolution VIIRS image to the DMSP data, we resample the raster to the DMSP resolution and then extract radiances of each pixel at the locations of the 10% sample that we have been using thus far. Naturally, there are considerable differences in the scale since the VIIRS-DNB records radiance. Note that radiance is measured in nano watt per steradian per square centimeter ($10^{-9}Wcm^{-2}sr^{-1}$). The difference in scale is reflected in the summary statistics of the VIIRS data. The mean is 3.98, the standard deviation is 18.65, and the maximum is 6567.42. The spatial Gini is much higher using the VIIRS data than in the radiance-calibrated data (0.79 vs. 0.60-0.65) which is owed to their improved sensors and finer resolution. Nevertheless, the top tail of the light distribution essentially exhibits the same properties.

Figure F-1 shows the Zipf plot for the VIIRS data. The shape is nearly linear, even high up in the tail and displays less curvature than the corresponding plot for the radiance-calibrated data. This also suggests that the radiance calibration process introduces noise and understates the Paretian nature of night lights.

Figure F-2 replicates a variant of the first figure of the paper by comparing the (normalized) light intensities indicated by the corrected data in 1999 to those recorded by the VIIRS DNB instrument in 2015. We downsampled the VIIRS data to the same resolution as the DMSP data by averaging the higher resolution pixels. The figure shows that the city profiles for the two mature cities (New York and London) are similar, even though the instruments, underlying scales, overpass times, and reference years differ. For New Delhi and Johannesburg, there is some discrepancy in the relative brightness of

sub-centers which may be due to recent developments or fundamental differences in how the data is recorded. As expected, the corrected data is considerably smoother than the VIIRS data, as it is derived from the radiance-calibrated data, which has both a much wider ground footprint and is subject to blooming (overglow).

Table F-1 replicates the results of the rank regressions using the VIIRS data. The results are qualitatively similar to those obtained with the radiance-calibrated data, but some small differences are notable. In particular, the estimated shape parameters are a bit higher for top shares of 3% to 5% but then also appear to be more stable in the upper tail. Since the VIIRS data are from five years after the most recent radiance-calibrated image and have a different overpass time, it is difficult to identify the source of these discrepancies.

Figure F-1 – Zipf plot using the top 2% of pixels in the VIIRS data
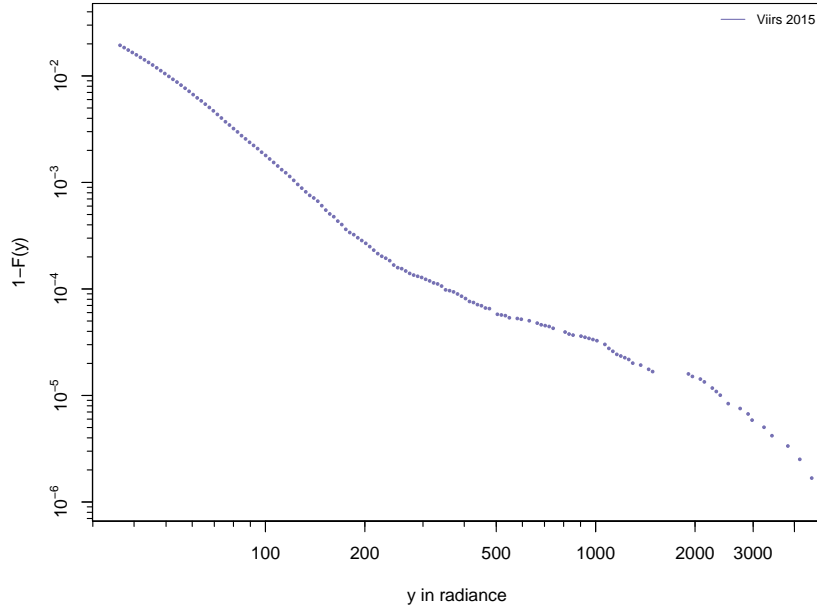
Notes: The figure shows a Zipf plot for the top 2% of all pixels of the VIIRS data, after resampling the data to the DMSP-OLS grid and resolution. The figure uses logarithmic binning to reduce noise and sampling errors in the right tail of the distribution (see Newman, 2005). There are about 140 bins in the tail, where the exact number depends on the range of the input data. The VIIRs pixels correspond to the same 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level obtained from Hsu et al. (2015) and used in the rest of the paper.

Figure F-2 – Corrected lights in 1999 versus VIIRS in 2015

(a) New York City
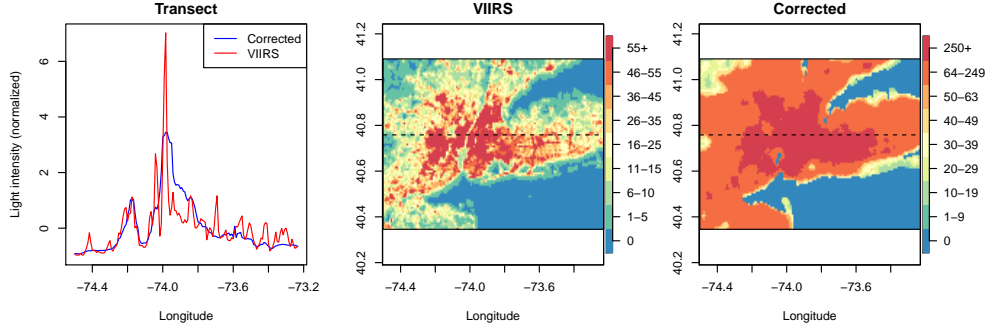
(b) London
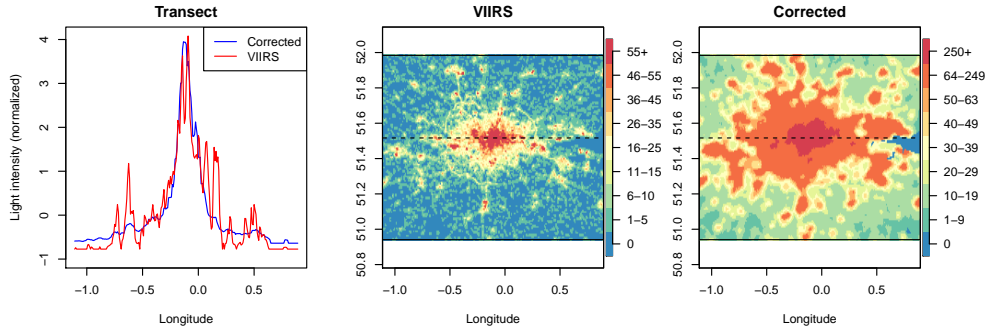
(c) New Delhi
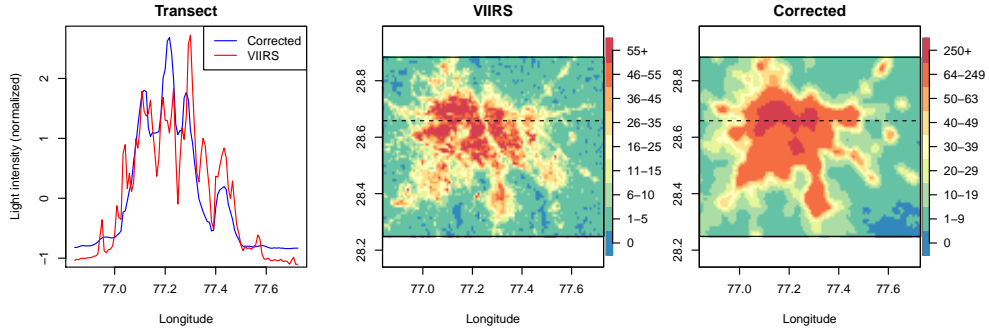
(d) Johannesburg
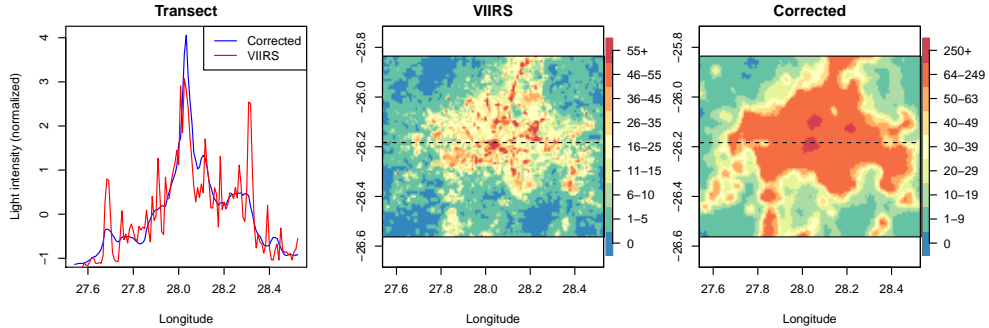
Notes: Comparison of the light intensities recorded our corrected lights in 1999 (as in Figure 1 in the paper) and by the VIIRS satellite in 2015 within four major cities. The left panel show the normalized light intensity along a longitudinal transect through the brightest pixel in each city. The middle panel shows a map based on the VIIRS data (in $10^{-9} W cm^{-2} sr^{-1}$). The right panel shows the same map using the corrected data (in DN). Both data have been binned and the color scales were adjusted so as to be visually comparable. Dashed lines indicate the transect path.

xxiv

Table F-1 – Rank regressions based on the VIIRS data in 2015

Unrestricted regressions			Hill estimates		Restricted regressions
			Panel a) Top 5%		
y_i	-1.9603	α	1.5876	α	1.7331
	(0.0013)		(0.0065)		(0.0100)
y_c	2.0405				
	(0.0015)				
R^2	0.9879				
Observations	59,633		59,633		59,633
			Panel b) Top 4%		
y_i	-2.0831	α	1.7331	α	1.8747
	(0.0013)		(0.0079)		(0.0121)
y_c	2.1479				
	(0.0015)				
R^2	0.9909				
Observations	47,705		47,705		47,705
			Panel c) Top 3%		
y_i	-2.2150	α	1.9144	α	2.0419
	(0.0014)		(0.0101)		(0.0153)
y_c	2.2622				
	0.0016				
R^2	0.9933				
Observations	35,780		35,780		35,780
			Panel b) Top 2%		
y_i	-2.3438	α	2.1671	α	2.2481
	(0.0017)		(0.0140)		(0.0206)
y_c	2.3665				
	(0.0019)				
R^2	0.9939				
Observations	23,854		23,854		23,854
			Panel e) Top 1%		
y_i	-2.3778	α	2.4716	α	2.4235
	(0.0033)		(0.0226)		(0.0314)
y_c	2.3682				
	(0.0036)				
R^2	0.9888				
Observations	11,927		11,927		11927

Notes: The table uses the VIIRS data to repeat three regressions which were conducted with the radiance-calibrated data before: the unrestricted OLS rank regression eq. (E-1) and the restricted regression eq. (E-2) using both the OLS and the Hill estimator. Standard errors are reported in paretheses. For the OLS restricted rank regression, these are the asymptotic standard errors computed as $(2/N)^{1/2}$. The data are a 10% representative sample of all non-zero lights in the radiance-calibrated data at the pixel level, where each pixel is 30×30 arc seconds.

G   An analytical top-coding correction

Researchers are often interested in aggregate measures, such as average luminosity or light inequality in a region or a country. Here we present simple formulas to correct these summary statistics for top-coding. These corrections work with arbitrary thresholds and Pareto shape parameters.

Mean luminosity: The top-coding corrected mean luminosity μ of a country or region is simply the weighted average of the bottom and top means μ_B and μ_T. If the latter is the mean of a Pareto distribution starting at y_c, we have

$$\mu = \omega_B \mu_B + (1 - \omega_B)\mu_T = \omega_B \mu_B + (1 - \omega_B)\frac{\alpha}{\alpha - 1}y_c \tag{G-1}$$

where ω_B and $\omega_T = 1 - \omega_B$ are the shares of pixels below and above the threshold. A simple numerical illustration shows how correcting for top-coding drives up the mean luminosity. If top-coding starts at $y_c = 55$, affects 5% of the study area of interest, α is 1.5 and mean luminosity in the non-top-coded pixels is $\mu_B = 10$, then the corrected mean luminosity is 17.75 rather than 12.25.

Spatial Gini coefficients: The overall Gini coefficient can be written as the weighted sum of the bottom-share and top-share Ginis (i.e., the within-group Gini) as well as the difference between the top share of total lights minus the top share of pixels (i.e., the between-group Gini), such that

$$G = \omega_B \phi_B G_B + \omega_T \phi_T G_T + [\phi_T - \omega_T], \tag{G-2}$$

where the shares of all light accruing to the top and bottom groups are $\phi_B = \omega_B \mu_B/\mu$ and $\phi_T = \omega_T \mu_T/\mu$, and $G_T = 1/(2\alpha - 1)$. A greater share of top-coded pixels ω_T, brighter top-coded pixels ϕ_T, and a greater spread in the distribution of the top-coded data G_T all increase the size of the correction.

The above decomposition of the Gini coefficient can be derived by defining the Gini coefficient over multiple groups as in Mookherjee and Shorrocks (1982)

$$G = \frac{1}{2N^2\mu}\sum_i \sum_j |y_i - y_j| \tag{G-3}$$

$$= \frac{1}{2N^2\mu}\sum_k \left(\sum_{i \in N_k}\sum_{j \in N_k} |y_i - y_j| + \sum_{i \in N_k}\sum_{j \notin N_k} |y_i - y_j| \right) \tag{G-4}$$

$$= \sum_k \left(\frac{N_k}{N}\right)^2 \frac{\mu_k}{\mu}G_k + \frac{1}{2N^2\mu}\sum_k \sum_{i \in N_k}\sum_{j \notin N_k} |y_i - y_j|. \tag{G-5}$$

where G_K is the within group Gini coefficient of group k. The second term is a measure of group overlap including their between group differences.

Perfect separation (no overlap between groups) implies $\sum_{i \in N_k} \sum_{j \in N_h} |y_i - y_j| = N_k N_h |\mu_k - \mu_h|$. Hence, we can simplify equation eq. (G-5) to

$$G = \sum_k \left(\frac{N_k}{N}\right)^2 \frac{\mu_k}{\mu} G_k + \sum_k \sum_h \frac{N_k N_h}{2N^2 \mu} |\mu_k - \mu_h| . \tag{G-6}$$

With two bottom and top groups $k, h \in \{B, T\}$ (where $\mu_T > \mu_B$) and some algebra, this becomes

$$G = \left(\frac{N_B}{N}\right)^2 \frac{\mu_B}{\mu} G_B + \left(\frac{N_T}{N}\right)^2 \frac{\mu_T}{\mu} G_T + \left[\left(\frac{N_T}{N}\right)^2 \frac{\mu_T}{\mu} - \frac{N_T}{N} \right] . \tag{G-7}$$

Now define the pixel shares below and above the threshold as ω_B and ω_T, where $\omega_T = 1 - \omega_B$ and the group's share of all income (light) as $\phi_B = \omega_B \frac{\mu_B}{\mu}$ and $\phi_T = \omega_T \frac{\mu_T}{\mu}$ to obtain eq. (G-2) above.

H Characteristics of the corrected data

In this section we compare the back-on-the-envelop analytical corrections from the previous section with our corrected data at the pixel level, examine the correlations between our corrected data and the radiance calibrated data, and discuss the size of the top-coding correction around the world.
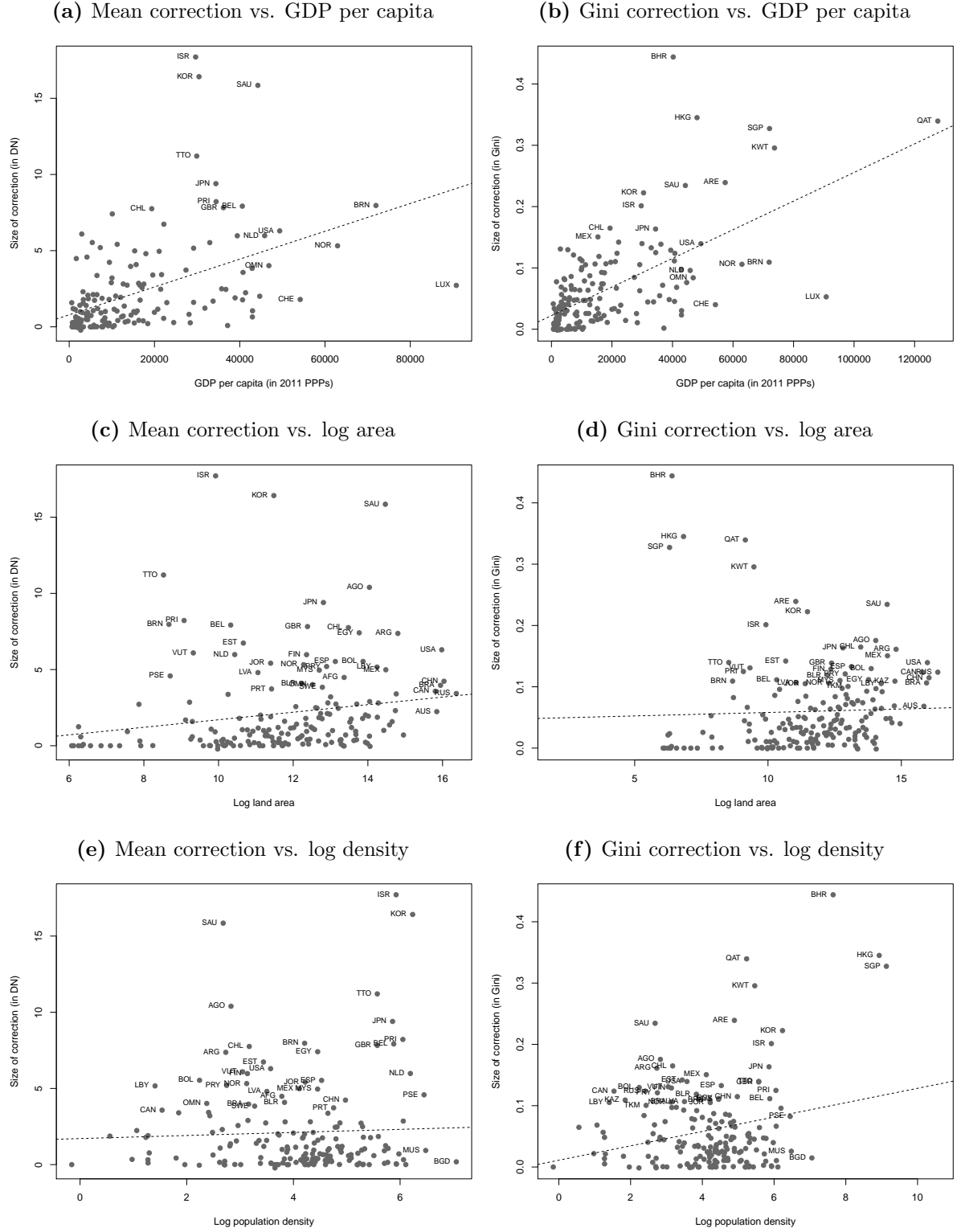
Figure H-1 illustrates the size of the correction in different countries with various scatter plots. As expected, the same characteristics that drive the number of top-coded pixels (see Section 2 in the paper) turn out to be predictive of the size of the correction in terms of country-wide mean luminosity and inequality in light. The correction is strongly increasing in GDP per capita, weakly in country size and moderately in population density. Numerous developing countries experience sizable corrections (such as Egypt, Paraguay or Mexico). City states, such as Singapore, have large top-coding corrections, as do smaller countries, like Israel and Estonia. Nevertheless, even large countries like the US experience a sizable increase in both mean luminosity (plus 7 DN) and the Gini coefficient (plus 14 percentage points). No single factor captures all the relevant heterogeneity.

Figure H-2 plots the time series graph of global inequality in lights from 1992 to 2013, both before and after the top-coding correction based on eq. (G-2). Parameter values of 1.4 and 1.6 serve as comparison bands for the benchmark case of 1.5. The global distribution of lights became slightly more unequal over the 1990s, remained flat in the first decade of the new millennium and then became temporarily more equal in the aftermath of the global financial crises and great recession. However, this year-to-year variation is completely swamped by the size of the top-coding correction.

Table H-1 reports mean luminosity and the Gini coefficient of inequality for 2010, using a wider range of parameters as robustness checks in the analytical correction in eq. (G-2). The parameters for this sensitivity analysis were chosen because of their equal distance to 1.5 rather than empirical relevance. Working with a smaller (larger) parameter than our benchmark $\alpha = 1.5$ implied more (less) inequality in the tail of the light distribution. The corrections are consequently larger (smaller). We can also see that parameter values of 1.4–1.6 only lead to very small differences in the magnitude of the correction. Also, using a higher α does not change the magnitude of the correction as much as using a smaller α, as the comparison of the extreme values of 1.2 and 1.8 shows.

Table H-2 reports the mean luminosities and global Gini coefficients before and after the correction for each satellite, using both the analytic formulas and the data corrected at the pixel level. Our geo-referenced pixel-level replacement comes close to the analytic solutions but is generally more conservative (due to the fixed upper bound). Mean luminosity increases on average from 12.7 DN to 15.3–16.6 DN and inequality in lights from 0.47 to 0.56–0.59.

Figure H-1 – Size of the correction and country characteristics

(a) Mean correction vs. GDP per capita

(b) Gini correction vs. GDP per capita

(c) Mean correction vs. log area

(d) Gini correction vs. log area

(e) Mean correction vs. log density

(f) Gini correction vs. log density

Notes: Illustration of how the correction of means and Gini coefficients, based on eq. (G-1) and eq. (G-2), correlate with log GDP per capita (PPP), log land area, and log population density. The data are a 10% representative sample of all non-zero lights in satellite F182010. GDP and population data are from the World Development Indicators. For display purposes, the left panels exclude countries with a correction of mean luminosity larger than 20 DN, these are Singapore, Hong Kong, Qatar, Bahrain, Kuwait, and the UAE. These countries are included in all panels on the right.

xxix

Figure H-2 – Global Gini coefficient in lights before and after the correction
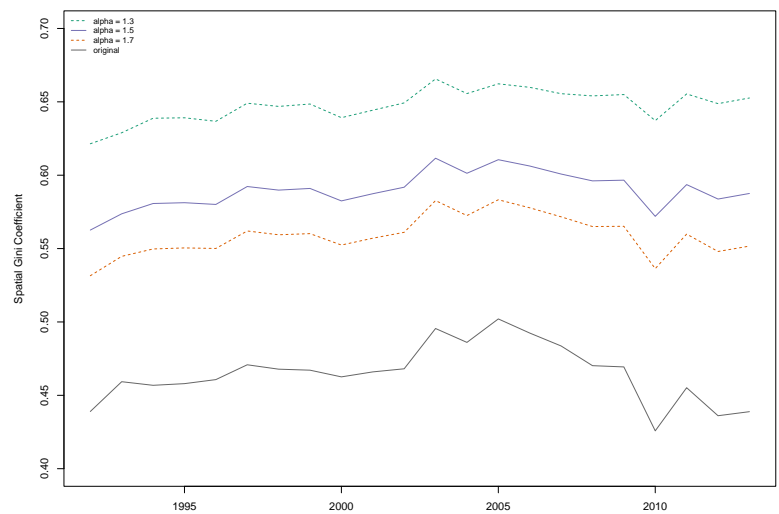
Notes: Illustration of the global top-coding correction. The figure shows global inequality in lights calculated by eq. (G-2) using the specified Pareto shape parameters. The input data are a representative 10% sample of non-zero lights. For years when more than two satellites flew concurrently, the values were averaged.

Table H-1 – Correction of global mean and Gini coefficient in 2010, different parameters

	Unadj.	*Pareto parameter $\alpha =$*						
		1.2	1.3	1.4	**1.5**	1.6	1.7	1.8
Mean luminosity	17.55	33.49	28.07	25.36	23.73	22.65	21.88	21.30
Spatial Gini	0.4258	0.6954	0.6372	0.5990	0.5720	0.5519	0.5363	0.5240
Top share (light)	0.2033	0.5825	0.5019	0.4487	0.4109	0.3828	0.3609	0.3435

Notes: The table computes the top-coding corrected mean and Gini coefficient of global inequality in lights for the year 2010 with different α parameters based on eq. (G-1) and eq. (G-2) with $y_c = 55$. The input data are a representative 10% sample of non-zero lights from satellite F18 in 2010.

Table H-2 – Satellite level statistics of the top-coding correction

Satellite	Top share (pixels)	Top share (light)		Mean luminosity			Gini coefficient		
		Unadj	Adj	Unadj	Form Adj	Pixel Adj	Unadj	Form Adj	Pixel Adj
F101992	0.0381	0.1663	0.3096	13.83	17.81	16.70	0.4390	0.5626	0.5334
F101993	0.0312	0.1568	0.2938	11.96	15.23	14.28	0.4593	0.5737	0.5456
F101994	0.0349	0.1754	0.3207	12.02	15.67	14.60	0.4783	0.5980	0.5684
F121994	0.0420	0.1733	0.3176	14.65	19.04	17.74	0.4353	0.5634	0.5316
F121995	0.0376	0.1733	0.3174	13.09	17.02	15.85	0.4580	0.5813	0.5505
F121996	0.0351	0.1670	0.3068	12.69	16.37	15.25	0.4607	0.5801	0.5494
F121997	0.0394	0.1766	0.3210	13.45	17.57	16.31	0.4540	0.5799	0.5477
F121998	0.0418	0.1816	0.3276	13.89	18.25	16.90	0.4474	0.5774	0.5436
F121999	0.0467	0.1915	0.3412	14.74	19.62	18.08	0.4447	0.5802	0.5449
F141997	0.0316	0.1739	0.3169	10.98	14.29	13.28	0.4876	0.6047	0.5747
F141998	0.0305	0.1680	0.3067	10.94	14.13	13.13	0.4883	0.6023	0.5720
F141999	0.0278	0.1648	0.3011	10.15	13.06	12.13	0.4895	0.6017	0.5714
F142000	0.0318	0.1689	0.3062	11.34	14.67	13.59	0.4852	0.6003	0.5687
F142001	0.0350	0.1817	0.3276	11.64	15.30	14.16	0.4856	0.6069	0.5754
F142002	0.0377	0.1872	0.3375	12.14	16.08	14.90	0.4896	0.6126	0.5818
F142003	0.0382	0.1930	0.3409	11.96	15.96	14.65	0.4928	0.6177	0.5836
F152000	0.0370	0.1685	0.3063	13.25	17.13	15.89	0.4399	0.5647	0.5308
F152001	0.0354	0.1645	0.3011	12.93	16.64	15.46	0.4463	0.5679	0.5351
F152002	0.0372	0.1700	0.3085	13.18	17.08	15.82	0.4465	0.5710	0.5370
F152003	0.0270	0.1582	0.2894	10.28	13.11	12.17	0.4982	0.6055	0.5751
F152004	0.0276	0.1642	0.2979	10.08	12.97	12.00	0.5080	0.6163	0.5853
F152005	0.0279	0.1604	0.2953	10.44	13.36	12.43	0.5115	0.6171	0.5886
F152006	0.0293	0.1666	0.2988	10.56	13.63	12.55	0.5135	0.6217	0.5892
F152007	0.0279	0.1547	0.2844	10.74	13.68	12.69	0.5049	0.6099	0.5795
F162004	0.0340	0.1734	0.3129	11.82	15.38	14.23	0.4641	0.5863	0.5528
F162005	0.0285	0.1642	0.2993	10.44	13.43	12.46	0.4926	0.6040	0.5732
F162006	0.0348	0.1707	0.3041	12.26	15.91	14.61	0.4714	0.5908	0.5546
F162007	0.0403	0.1861	0.3302	13.05	17.28	15.86	0.4624	0.5916	0.5554
F162008	0.0395	0.1832	0.3285	12.97	17.11	15.78	0.4702	0.5961	0.5622
F162009	0.0417	0.1862	0.3356	13.50	17.87	16.54	0.4694	0.5966	0.5644
F182010	0.0591	0.2033	0.3614	17.55	23.73	21.89	0.4258	0.5720	0.5361
F182011	0.0494	0.2020	0.3595	14.78	19.95	18.41	0.4552	0.5936	0.5598
F182012	0.0576	0.2118	0.3734	16.44	22.45	20.68	0.4361	0.5838	0.5481
F182013	0.0578	0.2151	0.3800	16.23	22.28	20.55	0.4389	0.5876	0.5530
Average	**0.0374**	**0.1765**	**0.3194**	**12.65**	**16.56**	**15.34**	**0.4691**	**0.5917**	**0.5595**

Notes: The table reports summary statistics of the global lights data before the top-coding correction and after the analytical, formula-based correction at the aggregate level (eq. (G-1) and eq. (G-2)) as well as the pixel-level correction from the paper. Column 1 reports the share of pixels above 55 DN, Column 2 and 3 the share of lights emitted by these top pixels respectively in the unadjusted and adjusted data set. Columns 4-6 and 7-9 report the mean luminosity and Gini coefficient, respectively for the unadjusted data, the analytical, formula-based correction and the pixel-level corrected data. All corrections use $\alpha = 1.5$ and $y_c = 55$ for the Pareto tail.

I Benchmarking exercises

I.1 Light-output elasticities at the national level

To validate our corrected data, we estimate light-output elasticities at the national level in the spirit of Henderson et al. (2012). Henderson et al. (2012) focus on the predictive relationship, that is, how lights predict changes in GDP, by running fixed effects regressions of log GDP on log lights per square kilometer. They report an elasticity income with respect to lights around 0.28. We then also examine the structural relationship where lights are on the left hand side and GDP on the right hand side (as in Chen and Nordhaus, 2011, Hu and Yao, 2019). Hu and Yao (2019) recently showed that there is significant non-linearity in the light production function which could be related to top-coding, as it appears to be weaker in the new VIIRS data (which is not top-coded). We replicate these results using a matched sample of the stable lights data, our top-coding corrected lights and the radiance-calibrated data for the seven years which all three data sources have in common over the period from 1996 to 2010. Note that both ways of estimating the relationship delivers biased estimates that depend differently on measurement errors in GDP and lights, and we make no attempt to correct for these biases here.

Table I-1 reports the estimates of the predictive relationship and shows that—even at the highly aggregated country level—our top-coding correction leads to marginal improvements. The corrected data always yield the highest within-R^2 and marginally larger estimates. When including quadratic terms on the right hand side, their estimated coefficients are smallest for the corrected and radiance-calibrated data. This suggests that top-coding plays some small role in the non-linearity of the predictive relationship.

Table I-2 turns things around and focuses on what is generally considered the structural relationship, that is, how lights react to changes in GDP. Here, too, we obtain the highest elasticity when we consider the corrected data and, by definition, exactly the same R^2 as in the reverse regression. In line with Hu and Yao (2019), we find evidence suggesting that there is non-linearity in the light production function. While it is weaker in the corrected data, it does not appear to qualitatively differ across the three data sources. Of course, these estimates are subject to measurement errors in GDP.

In summary, the results are not materially different at the national level, so that for an analysis of the light-output relationship at such a high level of aggregation either data can be used without explicitly considering the role of top-coding. Note that it is not clear whether we should expect the corrected data to deliver regression coefficients which are closer to the radiance-calibrated data. The spectral mixing process of Hsu et al. (2015) created a lot of noise in areas which are not top-coded and provide the overwhelming share of the variation analyzed by these regressions, so it is not implausible that our corrected data delivers the strongest results.

Table I-1 – Light-output elasticity, country-level, 1996–2010

	(1) Stable	(2) Corrected	(3) Radiance	(4) Stable	(5) Corrected	(6) Radiance
	Dependent variable: GDP in 2005 PPPs					
$\ln \text{LIGHTS}_{it}$	0.275***	0.278***	0.233***	0.193***	0.221***	0.182***
	(0.067)	(0.064)	(0.054)	(0.047)	(0.046)	(0.034)
$(\ln \text{LIGHTS}_{it})^2$				-0.020***	-0.015**	-0.017*
				(0.007)	(0.006)	(0.010)
Country FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Within-R^2	0.172	0.184	0.147	0.202	0.207	0.170
Observations	1288	1288	1288	1288	1288	1288
Countries	186	186	186	186	186	186

Note(s): The table reports panel FE estimates of the predictive relationship between light intensity and GDP within countries. $\ln \text{LIGHTS}_{it}$ are defined as the log average light intensity per sq. km. $\ln GDP_{it}$ is the log of GDP in 2005 PPPs from the World Development Indicators. The specifications are variants of $\ln \text{GDP}_{it} = \beta \ln \text{LIGHTS}_{it} + \gamma(\ln \text{LIGHTS}_{it})^2 + c_i + s_t + \epsilon_{it}$ where c_i is a country fixed effect, and s_t are year dummies. Country-clustered standard errors are provided in parentheses. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table I-2 – Light production function, country-level, 1996–2010

	(1) Stable	(2) Corrected	(3) Radiance	(4) Stable	(5) Corrected	(6) Radiance
	Dependent variable: Log lights per sq. km					
$\ln \text{GDP}_{it}$	0.628***	0.663***	0.632***	0.586***	0.627***	0.589***
	(0.101)	(0.093)	(0.076)	(0.080)	(0.077)	(0.073)
$(\ln \text{GDP}_{it})^2$				-0.037***	-0.032***	-0.038***
				(0.010)	(0.010)	(0.010)
Country FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Within-R^2	0.172	0.184	0.147	0.200	0.204	0.172
Observations	1288	1288	1288	1288	1288	1288
Countries	186	186	186	186	186	186

Note(s): The table reports panel FE estimates of the light production function within countries. $\ln \text{LIGHTS}_{it}$ are defined as the log average light intensity per sq. km. $\ln GDP_{it}$ is the log of GDP in 2005 PPPs from the World Development Indicators. The specifications are variants of $\ln Lights_{it} = \beta \ln \text{GDP}_{it} + \gamma(\ln \text{GDP}_{it})^2 + c_i + s_t + \epsilon_{it}$ where c_i is a country fixed effect, and s_t are year dummies. Country-clustered standard errors are provided in parentheses. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I.2 Light-output elasticities at the subnational level

Lights are particularly useful as a proxy for economic activity at the regional level or other smaller geographies. This is precisely where the influence of top-coding will be more pronounced. We now study how top-coding affects the non-linearity of the estimated light production function (in the spirit of Hu and Yao, 2019) in U.S. counties and German districts. Both countries publish subnational accounts data of the highest quality, which minimizes the attenuation bias caused by measurement errors in GDP and allows us to directly study variation in the structural parameter of interest. While some developing countries publish similar data, they are likely to include sizable errors in the estimation of regional GDP, severely limiting what we can learn from them without addressing these biases.

Figure I-1 plots the raw data for 401 German districts (*Kreise*) from 2000 to 2013 and 3,080 U.S. counties from 2001 to 2013. It visualizes how much the top-coding correction changes the light-output relationship in regions that are moderately to very economically dense. Our correction appears to restore linearity at the top in Germany, while it appears to shift up the point at which the relationship starts to deviate from a line in the case of U.S. counties.

Figure I-2 plots the average light intensity per square kilometer indicated by both data sources over population density in German districts and U.S. counties. It illustrates that there is a similar quadratic relationship as in the light production function estimated above, which again disappears or becomes substantially weaker in the corrected data. As top-coding occurs primarily in dense urban areas, we take this as evidence that the stable lights data artificially create non-linearity in comparatively dense regions.

Table I-3 estimates the light production function for the German districts. We present two-way fixed effects estimates (columns 1 and 2), pooled estimates with year fixed effects (columns 3 and 4), and cross-sectional estimates (columns 5 and 6). In each case, we compare the estimates obtained for a quadratic in log GDP (per sq. km) for the stable lights and corrected data. The data are centered so that GDP and lights are zero at their sample averages. All columns using the stable lights data suggest that there is substantial non-linearity in the light production function, with decreasing elasticities of light w.r.t. income as economic density increases. We little evidence in favor of a negative and significant quadratic term when using the top-coding corrected data, no matter if we estimate this using only within district or also with between district variation. This suggests that top-coding seems to be driving a substantial portion of this non-linearity at the district level in Germany.

Table I-4 shows similar estimates for U.S. counties. Here, too, we find some evidence that the light production function is non-linear in economic density. As before, the estimates of the squared terms are negative and significant at the 10%-level or smaller

whenever the stable lights data are used. However, the non-linearity disappears when we consider changes in nighttime lights within counties and becomes considerably weaker (falls by 40%) when we also use between-county variation. Given that measurement errors in GDP should be small in this context and are likely to be unrelated to measurement errors in lights, we take this as additional evidence that top-coding creates artificial non-linearity at high economic densities by limiting how much the brightness of these regions can grow.

Figure I-1 – Non-linearity of light production function in Germany and USA

(a) Germany – Stable

(b) Germany – Corrected

(c) USA – Stable

(d) USA – Corrected
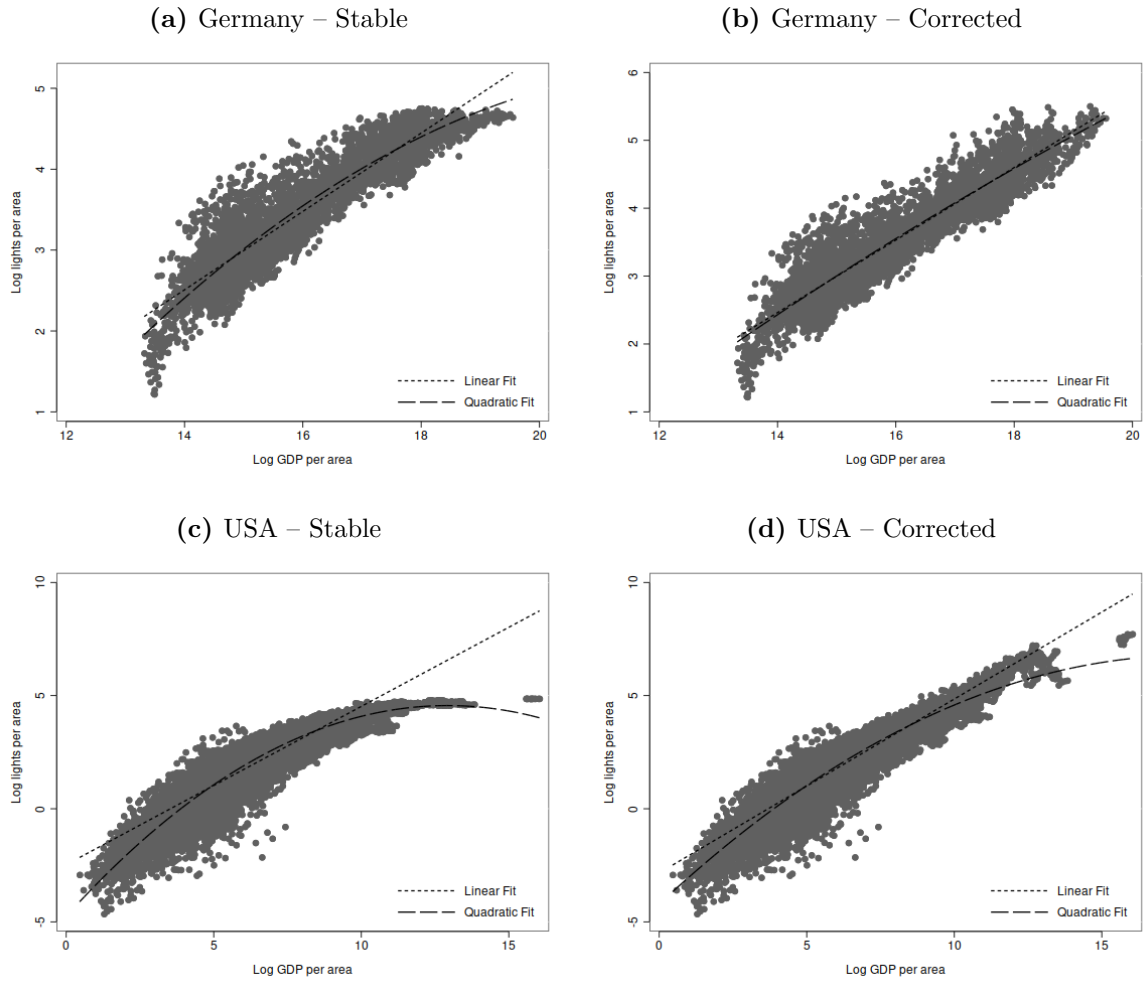
Notes: Illustration of the light-output relationship in German districts and U.S. counties. All panels show the log of average light intensity per sq. km. against the log of GDP per sq. km., including a linear and a quadratic fit. Panel (a) uses the stable lights data in German districts as the dependent variable. Panel (b) uses our corrected data in German districts as the dependent variable. Panel (c) uses the stable lights data in U.S. counties as the dependent variable. Panel (d) uses our corrected data in U.S. counties as the dependent variable.

Figure I-2 – Light density and population density in Germany and USA

(a) Germany – Stable

(b) Germany – Corrected

(c) USA – Stable

(d) USA – Corrected
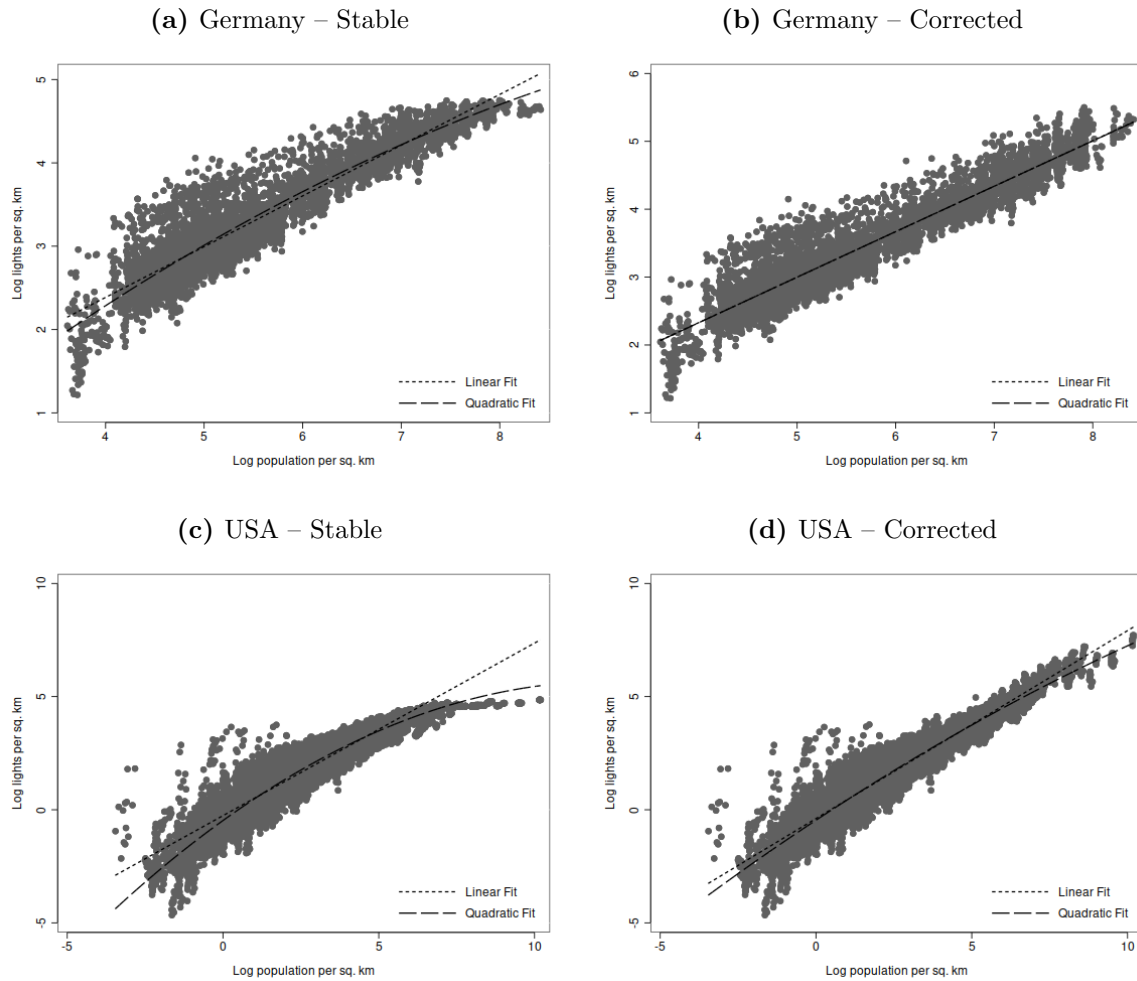
Notes: Illustration of the light-population relationship in German districts and U.S. counties. All panels show the log of average light intensity per sq. km. against the log of population density (population per sq. km.), including a linear and a quadratic fit. Panel (a) uses the stable lights data in German districts as the dependent variable. Panel (b) uses our corrected data in German districts as the dependent variable. Panel (c) uses the stable lights data in U.S. counties as the dependent variable. Panel (d) uses our corrected data in U.S. counties as the dependent variable.

Table I-3 – Estimated light production function, Germany at district level, 2000–2013

	Panel		*Pooled*		*Cross-section*	
	(1)	(2)	(3)	(4)	(5)	(6)
	Stable	Corrected	Stable	Corrected	Stable	Corrected
$\ln \text{GDP}_{it}$	0.222***	0.172***	0.510***	0.539***	0.502***	0.534***
	(0.034)	(0.029)	(0.008)	(0.009)	(0.009)	(0.010)
$(\ln \text{GDP}_{it})^2$	-0.076***	-0.005	-0.034***	-0.009*	-0.033***	-0.009
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
Time FE	✓	✓	✓	✓	–	–
District FE	✓	✓	–	–	–	–
Within-R^2	0.0671	0.00950	0.916	0.922	0.930	0.932
Observations	5614	5614	5614	5614	401	401
Districts	401	401	401	401	401	401

Notes: The table reports panel FE estimates of the light production function within and across German districts (Kreise). $\ln \text{LIGHTS}_{it}$ are defined as the log average light intensity per sq. km. $\ln GDP_{it}$ is the log of GDP per sq. km in current LCU from the GENESIS regional database. The specifications are variants of $\ln \text{LIGHTS}_{it} = \beta \ln \text{GDP}_{it} + \gamma (\ln \text{GDP}_{it})^2 + c_i + s_t + \epsilon_{it}$ where c_i is a district fixed effect, and s_t are year dummies. The cross-sectional regressions do not include any fixed effects. District-clustered standard errors are provided in parentheses. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table I-4 – Estimated light production function, USA at county level, 2001–2013

	Panel		*Pooled*		*Cross-section*	
	(1)	(2)	(3)	(4)	(5)	(6)
	Stable	Corrected	Stable	Corrected	Stable	Corrected
$\ln \text{GDP}_{it}$	0.375***	0.396***	0.750***	0.801***	0.751***	0.803***
	(0.039)	(0.038)	(0.005)	(0.005)	(0.005)	(0.005)
$(\ln \text{GDP}_{it})^2$	-0.006*	0.001	-0.056***	-0.034***	-0.058***	-0.035***
	(0.004)	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)
Time FE	✓	✓	✓	✓	–	–
County FE	✓	✓	–	–	–	–
Within-R^2	0.118	0.118	0.882	0.895	0.891	0.903
Observations	40039	40039	40039	40039	3080	3080
Counties	3080	3080	3080	3080	3080	3080

Notes: The table reports panel FE estimates of the light production function within and across U.S. counties. $\ln \text{LIGHTS}_{it}$ are defined as the log average light intensity per sq. km. $\ln GDP_{it}$ is the log of GDP per sq. km in current LCU from the Bureau of Economic Analysis (BEA) regional database (CAGDP1). The specifications are variants of $\ln \text{LIGHTS}_{it} = \beta \ln \text{GDP}_{it} + \gamma (\ln \text{GDP}_{it})^2 + c_i + s_t + \epsilon_{it}$ where c_i is a district fixed effect, and s_t are year dummies. The cross-sectional regressions do not include any fixed effects. District-clustered standard errors are provided in parentheses. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I.3 Light-wealth elasticities across cities within African countries

Lights are often used as a replacement for survey data on income and wealth at small geographies, when such survey data are unavailable or not widely available. Here we investigate whether top-coding makes a difference in the mapping of light to household wealth across African cities. As before, we are particularly interested in whether the relationship appears to be linear above the top-coding threshold or whether we observe a change in the relevant estimates. To this end, we use DHS survey data for urban sampling clusters in 29 African countries over the period from 1992 to 2013. Bruederle and Hodler (2018) compile this data to study correlations between various measures of household welfare and nighttime lights. Their primary measure is the DHS wealth index, which groups households in five groups according to scores derived from a principal component analysis of asset ownership. Bruederle and Hodler (2018) report a semi-elasticity of wealth to lights around 0.27.

Table I-5 reports estimates of the wealth-lights relationship across (up to) 7,601 urban survey clusters in 29 African countries. We present models that use within-country, within-country-year, and within-province-year variation to study how increases in the stable lights and corrected data translate into changes in the DHS wealth index. As survey locations change from survey to survey, we cannot exclusively use within-location variation but have to compare across locations within some geography (and year). In all panels, we allow the elasticity of wealth with respect to light to change for top-coded clusters. Since clusters consist of several pixels, we use the difference between the corrected light intensity and stable light intensity to identify top-coded clusters of varying severity. Following Michalopoulos and Papaioannou (2013), we always control for population density. This eliminates some of the comparability problems between household wealth and light emissions in the larger cluster. However, we cannot convert lights into a per household quantity without knowing the exact location and light emitted by the sampled household(s).

The results show that the wealth to light elasticity becomes very unstable in top-coded clusters when the stable lights data are used. In every specification, we find that the elasticity above the threshold is a multiple of the elasticity below the threshold and appears to increase in the threshold. This suggests that the variation in lights is limited above the threshold—recall that top-coding is now a matter of degree—although there is lots of variation in household wealth, so that small differences in average light intensities translate into large changes in household wealth. The results using the corrected data suggest that there is no such discontinuity for top-coded clusters. Instead, our estimates suggest that the relationship is linear in log lights, especially when we consider more severely top-coded clusters with a difference of at least 10 DN per sq. km. across both

data sources. Moreover, the within-R^2 is always moderately larger when the corrected data are used. Our results are similar whether we control for population density or not (not reported).

Table I-5 – Light-wealth relationship, urban DHS survey clusters, 1992–2013

	(1)	(2)	(3)	(4)	(5)	(6)
		Stable lights			Corrected lights	

Dependent variable: DHS wealth index

Panel a) Top-coded urban clusters when $\Delta DN_{it} > 0$

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \text{LIGHTS}_{ct}$	0.219***	0.221***	0.225***	0.230***	0.232***	0.230***
	(0.027)	(0.027)	(0.022)	(0.028)	(0.028)	(0.023)
$\ln \text{LIGHTS}_{ct} \times \text{TOP-CODED}_{ct}$	0.492**	0.485**	0.394	0.054**	0.053**	0.064**
	(0.225)	(0.228)	(0.232)	(0.024)	(0.025)	(0.031)
Within-R^2	0.261	0.263	0.177	0.275	0.278	0.187

Panel b) Top-coded urban clusters when $\Delta DN_{it} > 10$

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \text{LIGHTS}_{ct}$	0.230***	0.232***	0.236***	0.232***	0.234***	0.237***
	(0.026)	(0.026)	(0.023)	(0.027)	(0.027)	(0.024)
$\ln \text{LIGHTS}_{ct} \times \text{TOP-CODED}_{ct}$	1.683***	1.708***	1.957***	-0.002	-0.003	0.008
	(0.205)	(0.209)	(0.349)	(0.032)	(0.032)	(0.038)
Within-R^2	0.273	0.276	0.187	0.277	0.280	0.187

Panel c) Top-coded urban clusters when $\Delta DN_{it} > 20$

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \text{LIGHTS}_{it}$	0.233***	0.235***	0.238***	0.234***	0.236***	0.238***
	(0.026)	(0.027)	(0.023)	(0.027)	(0.027)	(0.025)
$\ln \text{LIGHTS}_{it} \times \text{TOP-CODED}_{it}$	1.593***	1.610***	1.964***	-0.043	-0.044	-0.028
	(0.439)	(0.442)	(0.458)	(0.037)	(0.036)	(0.041)
Within-R^2	0.276	0.279	0.186	0.278	0.281	0.187

	(1)	(2)	(3)	(4)	(5)	(6)
Pop. density	✓	✓	✓	✓	✓	✓
Country & Year FE	✓	–	–	✓	–	–
Country-Year FE	–	✓	–	–	✓	–
Province-Year FE	–	–	✓	–	–	✓
Observations	7601	7601	7405	7601	7601	7405
Countries	29	29	29	29	29	29

Note(s): The table reports regression estimates of the relationship between DHS wealth and light intensity within 2 km of urban DHS sampling clusters. $\ln \text{LIGHTS}_{it}$ are defined as the log average light intensity per sq. km. in a survey cluster using either the stable lights or corrected lights data. ΔDN_{it} is the difference between the corrected lights and stable lights per sq. km. in the cluster, so that positive values indicate the intensity of top-coding in the entire cluster. DHS_{it} is the average DHS wealth index in a survey cluster from Bruederle and Hodler (2018). The specifications are variants of $\text{DHS}_{it} = \beta \ln \text{LIGHTS}_{it} + \gamma \ln \text{LIGHTS}_{it} \times \text{TOP-CODED}_{it} + \theta \ln \text{POP. DENS.} + c_i + s_t + \epsilon_{it}$ where $\text{TOP-CODED}_{it} \equiv \mathbb{I}(\Delta DN_{it} > k)$ and k is a constant defined in the panel header, c_i is a country fixed effect, and s_t are year dummies. Country-clustered standard errors are provided in parentheses. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J Additional results for African cities

J.1 City growth in sub-Saharan Africa

Here we present additional results for the application of our correction to city growth in sub-Saharan Africa.

Figure J-1 shows the urban extents of selected cities and compares them with Google Earth images at the end of the periods we use to delineate urban areas, i.e., 12/1994 and 12/2013. The urban footprint detected by our algorithm coincides well with built-up structures (see Abrahams et al., 2018, for a more systematic comparison).

Figure J-2 compares the sum of lights in each of the three data series (stable lights, corrected data, radiance-calibrated data) for the available years. The radiance-calibrated data exhibit large fluctuations and jumps over time. On the other hand, the time series of the stable lights and corrected data fluctuate less and indicate a positive trend over the entire period.

Table J-1 reports descriptive statistics using only the years from 1996 to 2010, making the results comparable to the radiance-calibrated data. The stable lights data indicate similar growth rates for primary and secondary cities, just as in the longer sample. However, the corrected and radiance-calibrated lights suggest that primary cities grew faster at the intensive margin. The radiance-calibrated lights even indicate a negative average growth rate for secondary cities, which might be related to their large annual fluctuations. We observe positive growth in secondary cities for the corrected lights, but lower than that of primary cities.

Table J-2 gives an overview of the countries included in our study. The table reports the names of the primary city, the number of secondary cities, and the annualized growth rates for the stable lights and corrected data at the intensive margin. The corrections are larger in primate than in each country's secondary cities.

Table J-3 varies the minimum city size to account for the uncertainty in classifying cities as secondary. We focus on the strictest specification with country-year fixed effects. The main results are robust to larger thresholds and even increase when secondary cities are defined as fewer, larger cities. However, including smaller settlements increases noise, as their light intensity represents fewer and fewer data points per year, generating large jumps in their light intensity.

Table J-4 shows that our results are robust to excluding each of the African regions in turn. Using the corrected data, we reject the null that the interaction of the linear trend and the primacy dummy is zero in almost all regions. When we exclude East Africa, the coefficient marginally loses significance. Still, it remains within a standard error of the main result, while the stable lights data suggest that primate cities grow slower in this region.

Figure J-1 – Urban extents of selected cities

(a) Lagos, 12/1994
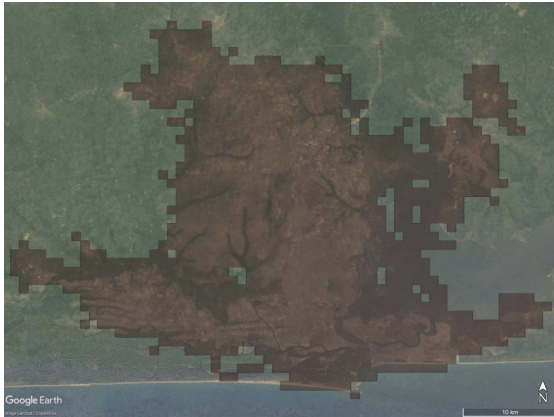
(b) Lagos, 12/2013

(c) Luanda, 12/1994

(d) Luanda, 12/2013

(e) Johannesburg, 12/1994

(f) Johannesburg, 12/2013
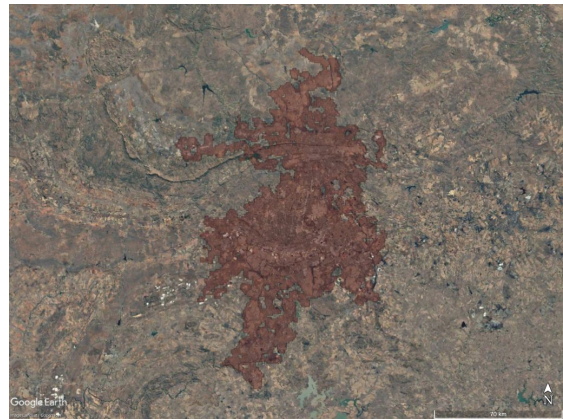
Notes: Illustration of the urban extents detection algorithm presented in the text. Note the differences in map scale. Comparison of 1992-1994 urban footprint with December 1994 and December 2013 Landsat/ Copernicus images obtained via Google Earth Pro. Google Earth images are used as part of their "fair use" policy. All rights to the underlying maps belong to Google.

Figure J-2 – Trends in light emitted by African cities, initial boundaries

(a) Primary cities   **(b)** Secondary cities
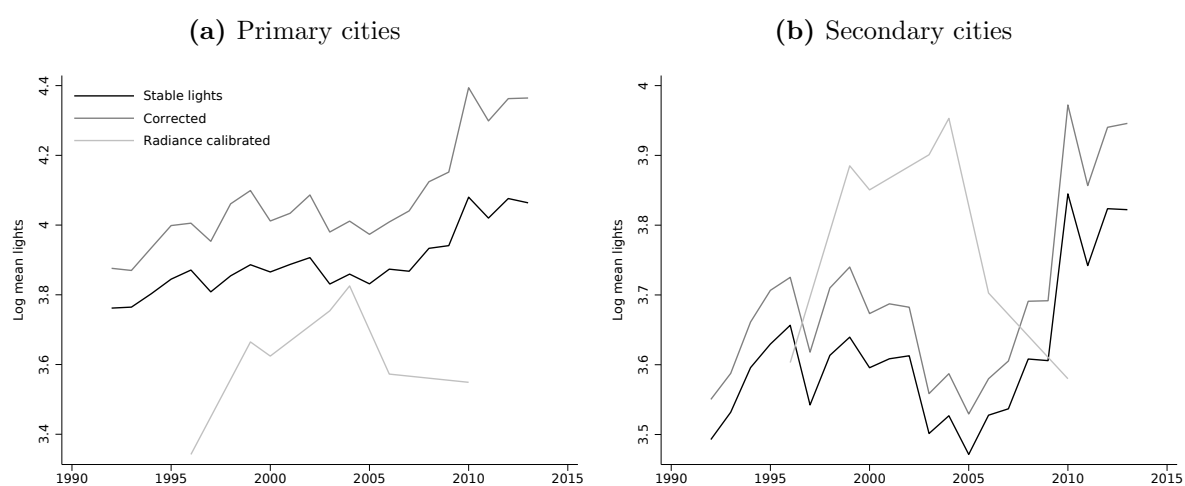
Notes: The figure illustrates the average light intensity over time in the sample of African cities (within the initial boundaries).

Table J-1 – Summary statistics for African cities, 1996–2010, envelopes and initial boundaries

	Stable Lights		Corrected Lights		Radcal Lights	
	Primate	Secondary	Primate	Secondary	Primate	Secondary
Panel a) Sum of lights in envelopes						
Sum in 1996	25227.59	2407.13	28488.13	2556.14	37816.48	2831.96
	(77801.08)	(7679.08)	(92532.01)	(9001.96)	(131460.00)	(12535.54)
Sum in 2010	35810.66	3169.00	46449.94	3515.06	46636.69	2934.17
	(98116.36)	(9094.25)	(135162.59)	(12027.36)	(144064.49)	(12468.45)
Annualized growth rate	0.0250	0.0196	0.0349	0.0228	0.0150	0.0025
Panel b) Sum of lights in initial boundaries						
Sum in 1996	22665.46	2095.41	25926.01	2244.42	13356.53	1986.41
	(73810.12)	(7378.00)	(88538.22)	(8720.04)	(20672.99)	(6679.74)
Sum in 2010	27934.93	2529.58	38238.95	2873.30	16425.64	1940.90
	(88793.92)	(8430.63)	(125218.42)	(11441.46)	(24210.20)	(6697.16)
Annualized growth rate	0.0149	0.0135	0.0278	0.0176	0.0148	-0.0017

Notes: The table reports a selection of summary statistics for African cities based on their envelopes (panel A) and their initial boundaries (panel B). In contrast to Table 3, the time period is from 1996 to 2010 to match the availability of the radiance-calibrated lights. Standard deviations in parentheses. Annualized growth rates are computed as $\frac{1}{14}(\ln x_{2010} - \ln x_{1996})$, where x refers to the data per group reported in the table.

Table J-2 – Annualized growth rates of cities in Africa, 1992–2013, intensive margin

Primate city	Primate city growth: Stable lights	Corrected	Secondary cities	Secondary city growth: Stable lights	Corrected	Country
Luanda	0.0282	0.0834	5	0.2034	0.2115	Angola
Bujumbura	0.0291	0.0317	3	0.1827	0.1823	Burundi
Cotonou	0.0328	0.0406	10	0.1775	0.1776	Benin
Ouagadougou	0.0265	0.0325	4	0.1842	0.1846	Burkina Faso
Gaborone	0.0114	0.0123	0			Botswana
Bangui	0.0193	0.0250	8	0.1642	0.1671	Central African Rep.
Abidjan	-0.0076	-0.0076	0			Cote d'Ivoire
Douala	0.0300	0.0324	3	0.1560	0.1560	Cameroon
Brazzaville	0.0124	0.0217	2	0.1897	0.1923	Congo (Dem. Rep.)
Kinshasa	0.0074	0.0181	12	0.1667	0.1679	Congo
Djibouti	0.0128	0.0229	18	0.1817	0.1819	Djibouti
Asmara	0.0205	0.0247	0			Eritrea
Addisabbeba	0.0217	0.0221	1	0.0627	0.0627	Ethiopia
Libreville	0.0227	0.0248	4	0.1771	0.1771	Gabon
Accra	0.0138	0.0182	4	0.1700	0.1718	Ghana
Conakry	0.0300	0.0314	0			Guinea
Banjul	0.0217	0.0284	18	0.1688	0.1694	Gambia
Bissau	0.0268	0.0273	2	0.1212	0.1212	Guinea-Bissau
Nairobi	0.0154	0.0154	0			Kenya
Maseru	0.0194	0.0213	12	0.1517	0.1525	Lesotho
Antananarivo	0.0390	0.0409	0			Madagascar
Bamako	0.0288	0.0309	5	0.1569	0.1569	Mali
Maputo	0.0167	0.0206	5	0.1526	0.1537	Mozambique
Nouakchott	0.0233	0.0303	1	0.1787	0.1787	Mauritania
Blantyre	0.0294	0.0365	2	0.1928	0.1925	Malawi
Windhoek	0.0341	0.0475	11	0.1870	0.1876	Namibia
Niamey	0.0130	0.0198	14	0.1735	0.1744	Niger
Lagos	0.0188	0.0181	5	0.1767	0.1767	Nigeria
Kigali	0.0175	0.0244	65	0.1643	0.1655	Rwanda
Alkhartum	0.0212	0.0218	1	0.1366	0.1366	Sudan
Dakar	0.0213	0.0329	11	0.1765	0.1776	Senegal
Freetown	0.0270	0.0270	0			Sierra Leone
Mogadishu	0.0635	0.0635	0			Somalia
Mbabane	0.0100	0.0186	209	0.1700	0.1715	Swaziland
Ndjamena	0.0210	0.0317	20	0.1873	0.1879	Chad
Sokode	0.0288	0.0303	5	0.1393	0.1393	Togo
Daressalaam	0.0231	0.0245	17	0.1618	0.1618	Tanzania
Kampala	0.0200	0.0228	2	0.1429	0.1429	Uganda
Johannesburg	0.0318	0.0367	4	0.1528	0.1528	South Africa
Lusaka	0.0185	0.0293	19	0.1769	0.1776	Zambia
Harare	0.0011	0.0014	19	0.1403	0.1403	Zimbabwe

Notes: The table reports summary statistics for the primary city and secondary cities in each country. The annualized growth rates are based on average light intensity and computed as $\frac{1}{21}(\ln x_{2013} - \ln x_{1992})$. The growth rate of secondary cities is an average across all such cities in the country.

Table J-3 – Growth regressions varying the minimum secondary city size, intensive margin

	Dependent variable: Log lights in the initial footprint							
	Stable lights	Corrected	Stable lights	Corrected	Stable lights	Corrected	Stable lights	Corrected
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Primate × trend	-0.217	0.351	-0.064	0.514**	0.294	0.895***	0.018	0.658***
	(0.267)	(0.234)	(0.266)	(0.228)**	(0.275)	(0.249)***	(0.238)	(0.209)***
	[0.269]	[0.233]	[0.269]	[0.228]**	[0.272]	[0.242]***	[0.225]	[0.193]***
Minimum size (in km^2)	6	6	9	9	15	15	25	25
City FE	✓	✓	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	14916	14916	12364	12364	8734	8734	5962	5962
Cities	678	678	562	562	397	397	271	271

Notes: The table reports the results of city-level panel regressions using the stable lights and top-coding corrected data with varying minimum size thresholds for secondary cities. All coefficients are scaled by 100 for readability. The specifications are variants of $\ln \text{LIGHTS}_{ijt} = \beta_1 t + \beta_2 (t \times P_{ij}) + c_{ij} + s_{jt} + \epsilon_{ijt}$ where t is a linear time trend, P_{ij} is an indicator for primate cities, c_{ij} is a city fixed effect and s_{jt} contains country-year fixed effects. Standard errors clustered at the city level are reported in parentheses. Conley errors with a spatial cutoff of 1,000 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets. Significant at: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table J-4 – Growth regressions excluding major regions, intensive margin

Dependent variable: Log lights in the initial footprint

	Stable lights (1)	Corrected (2)	Stable lights (3)	Corrected (4)	Stable lights (5)	Corrected (6)	Stable lights (7)	Corrected (8)
Primate × trend	0.110	0.724	0.141	0.568	-0.599	0.114	-0.080	0.515
	(0.356)	(0.297)**	(0.274)	(0.249)**	(0.248)**	(0.203)	(0.306)	(0.262)*
	[0.355]	[0.293]**	[0.278]	[0.251]**	[0.256]**	[0.205]	[0.301]	[0.256]**
Excluded region	West Africa		Middle Africa		East Africa		Southern Africa	
City FE	✓	✓	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	9240	9240	11528	11528	9900	9900	7018	7018
Cities	420	420	524	524	450	450	319	319

Notes: The table reports the results of city-level panel regressions using the stable lights and top-coding corrected data omitting major regions of Sub-Saharan Africa. All coefficients are scaled by 100 for readability. The specifications are variants of $\ln \text{LIGHTS}_{ijt} = \beta_1 t + \beta_2 (t \times P_{ij}) + c_{ij} + s_{jt} + \epsilon_{ijt}$ where t is a linear time trend, P_{ij} is an indicator for primate cities, c_{ij} is a city fixed effect and s_{jt} contains country-year fixed effects. Standard errors clustered at the city level are reported in parentheses. Conley errors with a spatial cutoff of 1,000 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J.2 City structure in Africa

In this section, we conduct an exploratory study of how the internal structure of cities is changing over time. Our aim is to show that nighttime lights can be used to better understand whether neighborhoods within African cities are becoming better connected or whether they increasingly resemble loose clusters of disconnected informal settlements. For this part of the analysis, we focus on the "envelope" of the city, that is, the maximum urban extent observed in both the initial and final boundaries. Focusing on the maximum urban extent allows cities to sprawl and become less connected over time (Harari, 2020), while cities in which slums are replaced with formal housing can increase in density and compactness.[6]

We compute two proxies for the variation of urban population density or within-city fragmentation, both of which are known from the literature on urban forms (e.g. see Tsai, 2005). Our first measure is the coefficient of variation of lights per km^2, a simple inequality measure capturing the variation of light intensities across an entire city. It is defined as the ratio of the standard deviation to the mean. A high (low) value indicates large (small) within-city differences in the dispersion of light. The index is not bounded from above.

Our second measure of fragmentation is Moran's I (Moran, 1950). Moran's I takes the precise location of each pixel within a city into account and indicates whether similar light intensities cluster together in space. It is defined as

$$I = \frac{N}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

where N is the number of pixels in the initial footprint of the city, w_{ij} are elements of an $N \times N$ inverse distance weight matrix, S_0 is the sum of all w_{ij}, x_i or x_j is the pixel-level light intensity, and $\bar{x}$ is mean luminosity.[7]

Positive values of Moran's I indicate that pixels are surrounded by others of similar luminosity or population density (positive autocorrelation), while negative values reflect a checkerboard pattern (negative autocorrelation). The index ranges from minus one to one. Light intensities within cities are positively spatially correlated but there is a clear ranking. The index continuously falls as we move from monocentric cities over polycentric cities to decentralized urban sprawl. A monocentric city in which luminosity slowly and gradually decreases from the densely populated center to the sparsely populated outskirts will have a higher Moran's I than a checkered city in which dense and sparsely populated

---

[6]We still focus on agglomerations which are now defined as all sub-cities which will eventually merge into a single metropolitan area. All results presented here are robust to using the initial footprint only.

[7]We work with a scaled version of Moran's I to make cities consisting of different numbers of pixels comparable, that is, we subtract its expected value under the null hypothesis of no spatial correlation: $I^* = I - \mathbb{E}[I] = I - (-1/(N-1))$.

areas take turns. We scale both indices by 100 for a better exposition.

Panel (a) of Figure J-3 illustrates the heterogeneity of urban structures on the subcontinent and shows that our light-based measures capture meaningful variation. Consider, for example, cities with a high Moran's I and relatively low coefficients of variation, such as Conakry, Dakar, and Cotounou. This combination indicates a regular structure with a bright center surrounded by similarly bright areas with a slow decay towards darker outskirts. Other cities with the same coefficient of variation have a much lower Moran's I. Their spatial distribution is considerably more fragmented, matching other accounts. A large part of Abidjan's population, for example, lives in slums characterized by illegal land tenure, buildings made out of non-permanent materials, and little or next to no infrastructure (UN-Habitat, 2003).

Johannesburg is an interesting case in terms of fragmentation. In 2000, it has one of the highest coefficients of variation and the lowest Moran's I in our sample of primate cities. Owing to a legacy of racial segregation during Apartheid, Johannesburg consists of alternating poor and rich neighborhoods which do not form a single integrated city. There is some limited evidence that this pattern may be changing. Panel (b) of Figure J-3 shows that we observe a moderate increase in Moran's I since the mid-2000s. The coefficient of variation is decreasing at the same time. This suggests that the different neighborhoods could be integrating, although the overall levels of inequality and fragmentation remain very high when compared with other cities in our sample.[8] Just as before, these raw time trends are only suggestive, as they include substantial measurement error.

We use the same methods to analyze these data in a more structured manner and focus on the differential between city types—i.e., we regress one of the measures of city fragmentation, F_{ijt}, on a linear time trend, an interaction of a linear time trend with an indicator for primate cities, P_{ij}, the log of lights per km^2 in the city, ln LIGHTS$_{ijt}$, and a set of fixed effects. We include the city-wide average light intensity to analyze their changing structure net of scale effects.

Table J-5 suggests two stylized patterns. First, we observe a decrease in the dispersion of lights over time which differs strongly across the two city types. Panel A shows that the coefficient of variation has been decreasing steadily over the period from 1992 to 2013. All three data sources agree on this trend. Panel B analyzes the development of Moran's I as a measure of spatial autocorrelation. The results suggest that there is no robust difference between primary and secondary cities using the corrected or radiance-calibrated data.

Table J-6 provides some preliminary evidence on whether urban form has an effect on city growth. To study this question, we regress log mean lights on the coefficient of variation or Moran's I in the previous year, a linear time trend, an interaction with primacy, and a combination of city and city-year fixed effects. Greater spatial inequality

---

[8]The evidence is stronger if we focus on the initial footprint and ignore that Johannesburg is sprawling. The rise towards the end of the sample is steeper and exceeds the values of the 1990s.

appears to have a robust negative effect on growth, but the impact of clustering is small and not robust. City structure and growth are endogenous, so that we only consider this an interesting partial correlation which could be explored in future research.

Our preferred interpretation of these findings is that Africa's biggest cities are at a crossroads. They are growing rapidly at the intensive and extensive margin, while the distribution of economic activity and people is starting to equalize across the city. Whether this conclusively demonstrates that poorer neighborhoods are becoming denser and brighter relative to the center but remain disconnected to other neighborhoods remains an open question. Further research might consider combining our corrected lights data with population grids and data on infrastructure to derive policy implications. Researchers interested in analyzing the structure of cities using the DMSP nighttime lights should also take the inherent auto-correlation in the data generating process into account (also see Abrahams et al., 2018, Gibson et al., 2021).

Figure J-3 – Varying structures of cities in Sub-Saharan Africa

(a) Fragmentation in primate cities, 2000   **(b)** Moran's I in Johannesburg
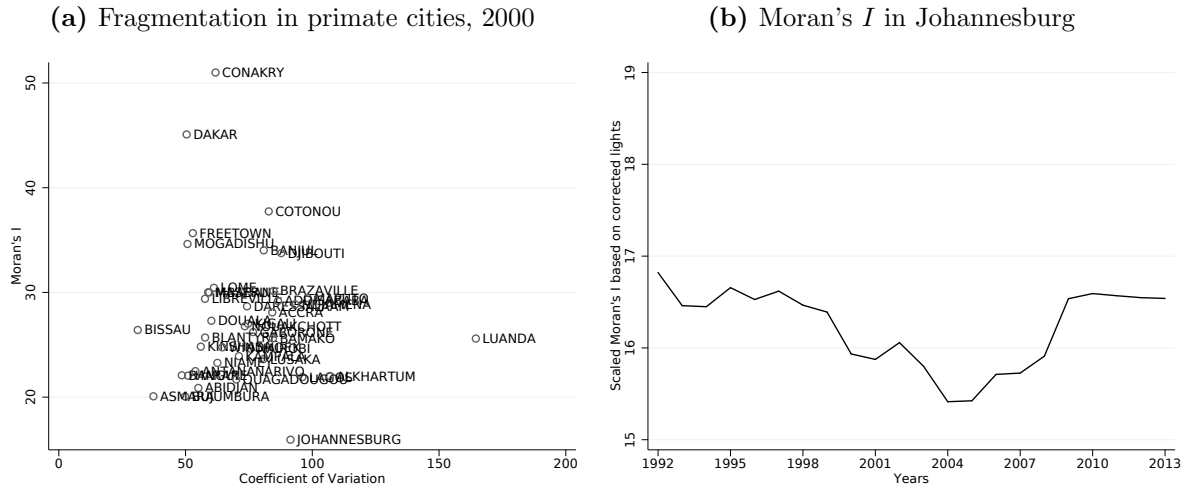
Notes: The figure illustrates urban structures in Sub-Saharan Africa. Panel (a) shows a cross-sectional scatter plot of the estimated coefficient of variation (CV) and Moran's I in 2000. Panel (b) displays the evolution of Moran's I in Johannesburg over the period from 1992 to 2013.

Table J-5 – Trends in fragmentation for African cities, envelopes

	Stable lights		*Corrected lights*		*Radcal lights*	
	Dependent variable: Varies by panel					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel a) Coefficient of variation in the envelope					
Primate × trend	-1.095	-1.048	-0.797	-0.657	-0.935	-0.992
	(0.146)***	(0.148)***	(0.128)***	(0.146)***	(0.217)***	(0.257)***
	[0.150]***	[0.146]***	[0.133]***	[0.145]***	[0.215]***	[0.252]***
Lights per km²	-19.477	-17.411	-18.732	-17.334	-5.630	-1.699
	(1.935)***	(1.978)***	(1.871)***	(1.997)***	(4.267)	(3.388)
	[2.096]***	[2.192]***	[2.028]***	[2.184]***	[4.349]	[3.776]
Observations	12356	12356	12356	12356	3932	3932
Cities	562	562	562	562	562	562
	Panel b) Moran's I in the envelope					
Primate × trend	-0.063	-0.051	-0.020	0.004	-0.092	-0.018
	(0.029)**	(0.040)	(0.027)	(0.036)	(0.052)*	(0.059)
	[0.030]**	[0.039]	[0.029]	[0.035]	[0.057]	[0.057]
Lights per km²	-1.191	-1.296	-1.232**	-1.456**	-0.902	-1.169
	(0.501)**	(0.618)**	(0.496)**	(0.617)**	(0.574)	(0.715)
	[0.509]**	[0.604]**	[0.504]**	[0.604]**	[0.585]	[0.691]
Observations	12356	12356	12356	12356	3932	3932
Cities	562	562	562	562	562	562
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	–	✓	–	✓	–
Country-Year FE	–	✓	–	✓	–	✓

Notes: The table reports results of city-level panel regressions using the stable lights, top-coding corrected data, as well as the radiance-calibrated data. The specifications are variants of $F_{ijt} = \beta_1 t + \beta_2(t \times P_{ij}) + \beta_3 \ln \text{LIGHTS}_{ijt} + c_{ij} + s_{jt} + \epsilon_{ijt}$, where F_{ijt} is either the coefficient of variation (Panel A) or Moran's I (Panel B), t is a linear time trend, P_{ij} is an indicator for primate cities, c_{ij} is a city fixed effect and s_{jt} contains are varying set of fixed effects (year, or country-year). Standard errors clustered at the city level are reported in parentheses. Conley errors with a spatial cutoff of 1,000 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table J-6 – Impact of fragmentation on city growth, envelopes

	Stable lights		Corrected lights		Radcal lights	
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Log lights in the envelope					
	Panel a) Impact of coefficient of variation on growth					
Primate × trend	0.003	0.001	0.007	0.006	0.029	0.032
	(0.002)	(0.002)	(0.003)**	(0.002)***	(0.008)***	(0.006)***
	[0.003]	[0.003]	[0.003]**	[0.002]***	[0.008]***	[0.005]***
Lagged CV	-0.006	-0.004	-0.006	-0.004	-0.003	-0.003
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
	[0.001]***	[0.001]***	[0.001]***	[0.001]***	[0.001]***	[0.001]***
Observations	11791	11791	11791	11791	1120	1122
Cities	562	562	562	562	560	562
	Panel b) Impact of Moran's I on growth					
Primate × trend	0.010	0.005	0.013	0.009	0.033	0.035
	(0.003)***	(0.003)**	(0.004)***	(0.002)***	(0.008)***	(0.006)***
	[0.003]***	[0.003]**	[0.004]***	[0.002]***	[0.008]***	[0.006]***
Lagged Moran's I	-0.003	-0.002	-0.003	-0.002	-0.002	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
	[0.002]	[0.002]	[0.002]	[0.002]	[0.003]	[0.003]
Observations	11791	11791	11791	11791	1120	1122
Cities	562	562	562	562	560	562
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	–	✓	–	✓	–
Country-Year FE	–	✓	–	✓	–	✓

Notes: The table reports the results of city-level panel regressions using the stable lights, top-coding corrected data, as well as the radiance-calibrated lights, where either the lagged coefficient of variation or Moran's I are used as regressors ($F_{ij,t-1}$). All coefficients are scaled by 100 for readability. The specifications are variants of ln $\text{LIGHTS}_{ijt} = \beta_1 t + \beta_2 (t \times P_{ij}) + \beta_3 F_{ij,t-1} + c_{ij} + s_{jt} + \epsilon_{ijt}$ where t is a linear time trend, P_{ij} is an indicator for primate cities, c_{ij} is a city fixed effect and s_{jt} contains a varying set of fixed effects (year, or country-year). Standard errors clustered at the city level are reported in parentheses. Conley errors with a spatial cutoff of 1,000 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets. Significant at: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Additional references

Abrahams, A., C. Oram, and N. Lozano-Gracia (2018). Deblurring DMSP nighttime lights: A new method using gaussian filters and frequencies of illumination. *Remote Sensing of Environment 210*, 242–258.

Bruederle, A. and R. Hodler (2018, 09). Nighttime lights as a proxy for human development at the local level. *PLOS ONE 13*(9), 1–22.

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.

Cirillo, P. (2013). Are your data really Pareto distributed? *Physica A: Statistical Mechanics and its Applications 392*(23), 5947–5962.

Coles, S. (2001). *An introduction to the statistical modeling of extreme values*. Springer Series in Statistics.

Eeckhout, J. (2009). Gibrat's law for (all) cities: Reply. *American Economic Review 99*(4), 1676–1683.

Elvidge, C. D., K. E. Baugh, J. B. Dietz, T. Bland, P. C. Sutton, and H. W. Kroehl (1999). Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sensing of Environment 68*(1), 77–88.

Elvidge, C. D., F.-C. Hsu, K. E. Baugh, and T. Ghosh (2014). National trends in satellite-observed lighting. In Q. Weng (Ed.), *Global Urban Monitoring and Assessment through Earth Observation*, Taylor & Francis Series in Remote Sensing Applications, Chapter 6, pp. 97–119. CRC Press.

Elvidge, C. D., D. Ziskin, K. E. Baugh, B. T. Tuttle, T. Ghosh, D. W. Pack, E. H. Erwin, and M. Zhizhin (2009). A fifteen year record of global natural gas flaring derived from satellite data. *Energies 2*(3), 595–622.

Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics 1*(1), 255–294.

Gabaix, X. and R. Ibragimov (2011). Rank - 1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics 29*(1), 24–39.

Gabaix, X. and Y. Ioannides (2004). The evolution of city size distribution. In J. V. Henderson and J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (1 ed.), Volume 4, Chapter 53, pp. 2341–2378. Elsevier.

Gibson, J., S. Olivia, and G. Boe-Gibson (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics 149*, 102602.

Harari, M. (2020). Cities in bad shape: Urban geometry in India. *American Economic Review 110*(8), 2377–2421.

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

Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics 3*(5), 1163–1174.

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

Hu, Y. and J. Yao (2019). Illuminating economic growth. IMF Working Paper no. 19/77, International Monetary Fund.

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

Mookherjee, D. and A. Shorrocks (1982). A decomposition analysis of the trend in UK income inequality. *Economic Journal 92*(368), 886–902.

Moran, A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika 37*(1/2), 17–23.

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics 46*(5), 323–351.

Nordhaus, W. and X. Chen (2015). A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics. *Journal of Economic Geography 15*(1), 217–246.

Tsai, Y.-H. (2005). Quantifying urban form: Compactness versus 'sprawl'. *Urban Studies 42*(1), 141–161.

Tuttle, B., S. Anderson, P. Sutton, C. Elvidge, and K. Baugh (2013). It used to be dark here. *Photogrammetric Engineering and Remote Sensing 79*(3), 287–297.

UN-Habitat (2003). *The challenge of slums: Global report on human settlements, 2003*. Earthscan Publications.

Ziskin, D., K. Baugh, F. C. Hsu, T. Ghosh, and C. Elvidge (2010). Methods used for the 2006 radiance lights. *Proceedings of the Asia-Pacific Advanced Network 30*, 131–142.