Journal of Urban Economics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Urban Economics

journal homepage: www.elsevier.com/locate/jue

Measuring the size and growth of cities using nighttime light☆

Rafael Ch[a], Diego A. Martin[b], Juan F. Vargas[c,*]

[a] New York University, Department of Politics, 19 West 4th Street, room 230, NY, NY 10003, USA
[b] Purdue University, Department of Economics, 425 W. State Street, Kenner Center, West Lafayette, IN 47907, USA
[c] Department of Economics, Universidad del Rosario, Calle 12C No. 4–69, Of. 315, Bogotá, Colombia

ARTICLE INFO

JEL classification:
O18
R12
R14

Keywords:
Size of metropolitan areas
Urbanization rates
Nighttime light data
Zipf's law
Gibrat's law

ABSTRACT

This paper uses nighttime luminosity to estimate a globally comparable measure of the size of metropolitan areas around the world for the years 2000 and 2010. We apply recently-proposed methodologies that correct the known problems of available nighttime luminosity data including blurring, instability of lit pixels overtime and the reduced comparability of night light images across satellites and across time. We then develop a protocol that isolates stable nighttime lit pixels that constitute urban footprint. When analyzed together with existing geo-referenced population datasets, our measure of the size of metropolitan areas can be used to compute urbanization rates, urban densities, and study the size distribution of cities. We show these applications and discuss how they compare with other available figures.

1. Introduction

The process of economic growth usually involves structural changes, of which urbanization is the most important one (Kuznets, 1968). In fact, for many social scientists urbanization is the hallmark of economic development. For example, in the absence of historical information on per capita income, several authors have used historical urbanization rates as a proxy of economic prosperity.[1] This implies that cities, as nuclei of the urbanization process, are the main drivers of economic growth.

But what is a city? Conceptually, cities -or metropolitan areas-[2] are the spatial integration of social and economic activity, and the higher the geographical concentration of households and firms, the greater the potential economic benefits net of congestion costs (Combes et al., 2019; Duranton and Puga, 2014; Rosenthal and Strange, 2004). However, in spite of our relatively consensual theoretical understanding of cities, its empirical measurement has proven more challenging. Empirical definitions usually rely on administrative borders, which do not necessarily correspond by the physical space where interactions between economic and social agents occur. This implies that stylized facts about patterns of urbanization across countries are likely to be biased because of measurement error (Roberts et al., 2016), which may partly explain, for instance, why the available cross-country estimates of the Zipf's law are

☆ We are grateful to Gilles Duranton, Cynthia Goytia, Pablo Sanguinetti and two anonymous referees for useful comments and discussion. We also thank seminar participants at CAF's RED 2017 workshop in Buenos Aires and the First Urban and Regional Economics Workshop at Universidad Javeriana. This paper was possible thanks to generous funding from CAF-Development Bank of Latin America.

* Corresponding author.
E-mail addresses: rafael.ch@nyu.edu (R. Ch), dmartinl@purdue.edu (D.A. Martin), juan.vargas@urosario.edu.co (J.F. Vargas).

[1] See, for example, Acemoglu et al. (2002) and Dittmar (2011). The historical urbanization data is estimated by Bairoch (1991).

[2] For simplicity, in this paper we refer to cities and metropolitan areas interchangeably. Our data-driven measure or urban agglomeration is arguably closer to what the latter are.

https://doi.org/10.1016/j.jue.2020.103254
Received 20 September 2018; Received in revised form 1 April 2020
Available online xxx
0094-1190/© 2020 Elsevier Inc. All rights reserved.

Please cite this article as: R. Ch, D.A. Martin and J.F. Vargas, Measuring the size and growth of cities using nighttime light, Journal of Urban Economics, https://doi.org/10.1016/j.jue.2020.103254

Table 1
Number of cities .

	2000		2010	
	Min (1)	Max (2)	Min (3)	Max (4)
East Asia and Pacific	1334	3000	1547	3569
Europe and Central Asia	1052	1397	980	1507
Latin America and Caribbean	768	842	738	987
Middle East and North Africa	241	287	257	352
North America	340	366	367	471
South Asia	361	612	325	786
Sub-Saharan Africa	130	192	136	232
Total	4248	6635	4656	7660

Notes. Columns report minimum and maximum number of cities estimated using *BEAM* and Landscan georeferenced population datasets, considering cites with more than 50,000 inhabitants and applying a pixel stability threshold of 0%, 20%, or 50%, and a splitting threshold of metropolitan areas of less than 1 km, 2 km, 3 km, 4 km or 5 km.

not robust.[3] As put by Duranton and Puga (2014) "The empirical validity of Zipf's -law is hotly debated" (p.833), and the cross-country evidence (Rosen and Resnick, 1980; Soo, 2005) has been interpreted both for and against the Zipf's law. Chauvin et al. (2017), for example, test the existence of the law in Brazil, China, India and the U.S., using respectively micro-regions, cities, districts and metropolitan areas. These are heterogeneous measures whose comparison the authors recognize as "debatable" (p.20).

In this paper, we introduce a novel methodology to translate the concept of city into a standard measurement of the size of a metropolitan area that is independent of country-specific administrative criteria of city limits, allowing for global comparisons. Specifically, following recent remote sensing and economic development literature -reviewed by Donaldson and Storeygard (2016)-, our measure leverages on the availability of satellite images of night lights.[4]

We adopt two main steps: first, nighttime light data are cleaned and processed; second, processed lit pixels are classified to identify the extension of cities. The first step recognizes that nighttime light data suffers from several sources of measurement error, which we correct by combining insights from a large body of literature on remote sensing (Abrahams et al., 2018; Elvidge et al., 2009; Imhoff et al., 1997; Liu et al., 2012).[5] Importantly, we do not claim any improvement on the remote sensing methodological literature. Rather, we provide a systematic and replicable algorithm on how to correct several sources of potential bias for the specific purpose of classifying lit pixels as urban areas.

The second step classifies metropolitan areas as spatial conglomerates of 1 km^2 nighttime *lit* pixels built around a pre-identified urban core. An important challenge is that some lit pixels are unstable. That is, they are detected by the satellite in only a fraction of the daily pictures. Clearly, imposing a stability threshold (that is, requiring pixels to be lit in a minimum fraction of the satellite images) affects the estimation of the size of metropolitan areas. Since we do not hold a strong prior about the minimum pixel stability, the paper presents estimations using various thresholds.[6]

As a result, we end up with a world-wide geo-referenced dataset of the location and area of a maximum of 6,635 (7,660) metropolitan areas of at least 50,000 inhabitants in 2000 (2010). The actual number of cities changes by varying the pixel stability threshold, as well as the threshold of the width of the urban corridor used to split very large metropolitan areas into underlying cities.[7] We produce a total of 15 different estimates of cities by varying these thresholds. Table 1 reports the minimum and maximum estimated number of cities for the entire world as well as each one of the seven world regions identified by the World Bank, both for 2000 and for 2010. Based on the Spanish acronym of this Metropolitan Areas Extension Database (*Base de Extensión de Áreas Metropolitanas*), we call our resulting data BEAM.[8,9]

Our globally comparable dataset can be further overlaid with existing geo-referenced population data in order to estimate urban population (and thus urbanization rates) and population growth, as well as compute city-wide densities. In fact, we compare the distribution of urban area, urban population and population density across world regions and document several robust stylized facts. In particular, while the distribution of urban population in cities larger than 50,000 people is relatively similar across regions, regardless of their level of development, average urban area is larger in developed countries, especially North America. This implies that population density is higher in developing countries. We also compute urbanization rates for the entire world -conditional on cities having at least 50,000 inhabitants- and study the cross-country association between urbanization rates and income levels. We find that the correlation is positive in each cross section, but not when we exploit the panel structure of BEAM and control for country and time fixed effects. In this case the correlation either disappears or may even become negative. This may well be explained by the specific sample period, which covers a decade of low global growth rates (2000 to 2010).

In addition to providing details about the computation of BEAM and describing these basic stylized facts, we illustrate the advantages of this measure studying the size distribution of cities in the entire world and compare our results to recent evidence provided by Chauvin et al. (2017) for selected developed and developing countries. For instance, when we test Zipf's law using BEAM, our estimates of the rank size rule are smaller than those of Chauvin et al. (2017) -with the exception of China- and smaller than what it is predicted by Zipf's law. This implies that, in three out of the four countries analyzed by these authors, the largest cities are larger than what Zipf's law would predict.[10]

[3] Country-specific measures of cities or metropolitan areas are based on arbitrary thresholds, many of which depend on the quality of local censuses and their minimum collection unit size. Since the latter depends on state and administrative capacity, country-specific measurement error is likely correlated with the level of development, and thus it is potentially non-classical.

[4] Nighttime light has been shown to be a reliable proxy for economic activity both nationally and in geographically small areas (Bleakley and Lin, 2012; Goldblatt Ran et al., 2018; Henderson et al., 2011; Lowe, 2014; Michalopoulos and Papaioannou, 2013; Storeygard, 2016; Weidmann and Schutte, 2016). This implies that it can be used to identify geographically integrated economic markets, thus linking our practical and conceptual definitions.

[5] In particular, we correct censoring problems using a simple clipping procedure, apply a deblurring filter in order to compute an unbiased area for each city/satellite-year, and adjust the instability of lit pixels overtime and between satellites due to differences in satellite-specific crossing times, sensors' degradation and geographic misalignment.

[6] According to Abrahams et al. (2018) a 20% pixel stability threshold is good enough to remove ephemeral lights including cars and other types of spurious bright sources.

[7] The resulting distribution of cities includes a few very large and highly urbanized settlements that comprise what other datasets identify as separate cities. This is the case, for instance, of the New York City-Washington D.C. corridor of the U.S. East Coast. In these cases, the underlying cities are usually joint by urban corridors that expand along the main highways (and therefore generate lit pixels for reasons other than the sole highway illumination). Our estimates are robust to splitting these huge metropolitan areas among the underlying cities using different thresholds of the width of the connecting urban corridor. This additional step is conceptually not necessary, but increases comparability with other geo-referenced databases of city size and shape.

[8] Conveniently, beam also means ray of light in English.

[9] An older version of BEAM, which did not include all the correction procedures described in Section 2, and that was only available for a subset of regions of the world, was used as an input for the 2017 *Report on Economics and Development*, the flagship report of CAF-Development Bank of Latin America, which funded this research.

[10] This is also what Chauvin et al. (2017) find for the U.S. and China.

The rest of the paper is organized as follows: Section 2 discusses several measures used in the literature to estimate the size of metropolitan areas, and describes the methodology, including data problems, the data cleaning procedure and the classification of lit pixels into urban conglomerates. It also includes a brief description of how population counts can be used as an ex-post estimation procedure. Section 3 compares BEAM with existing datasets and summarizes world and region-level stylized facts regarding the number of cities, the distribution of city areas, population density and urbanization rates, including a discussion about the relationship between urbanization and income per capita. Section 4 establishes some basic facts about the size distribution of cities using comparable data for the whole world. Section 5 provides a discussion about potential disadvantages to estimate the size of metropolitan areas, and Section 6 concludes.

2. Measuring the area and size of metropolitan areas

Increasing urbanization rates have highlighted the need to re-define urban areas and explore methodologies to accurately approximate the size of cities. As cities grow, political administrative boundaries become obsolete to define metropolitan areas in a way that is consistent with modern commuting and exchange patterns. While the U.S. introduction of Metropolitan Statistical Areas (MSAs) in the 1950s established a general benchmark followed by other high income countries, many developing countries still lack institutional definitions for metropolitan areas, and thus rely heavily on administrative boundaries and simple aggregation procedures that are often at odds with patterns of urbanization.

There are at least three broad approaches used to define the extent and size of cities. First, the use of administrative geographic units aggregated through various iterative approaches given different thresholds, for instance minimum population counts or minimum density levels. In some cases, administrative units are agglomerated if they are contiguous and economically "integrated", as in the case of Brazil, where contiguous municipalities are integrated into *microregions* conditional on having similar economic features, according to the Brazilian Institute for Geography and Statistics. However, non-comparability arises to the extent that countries define "integration" and "contiguity" differently, and use arbitrary thresholds. For instance, to the extent that relatively more rural areas are integrated into a *microregion*, that would lead to a potential over estimation of the size of the metropolitan area.

A second approach relies on commuting patterns, understanding metropolitan areas as integrated labor markets. In most developed countries, census or survey data track commuting patterns by asking respondents about the location of their residence and that of their job. Most studies that use this approach predefine metropolitan cores using ancillary population data, and then use commuting patterns to determine how to aggregate it. More recently, however, Duranton (2015) showed that there is no need to rely on pre-defined urban cores by proposing an algorithm to delineate metropolitan areas based solely on commuting thresholds. Nonetheless, reliable data on commuting patterns is not available for most countries, especially developing ones. Moreover, while commuting patterns portray the reach of labor markets, they leave aside input-output linkages which might cover wider areas, or knowledge spillovers which tend to revolve around shorter distances.[11]

Note that both the administrative unit agglomeration and commuting pattern approaches rely on arbitrary thresholds. While some studies show that results hold when modifying thresholds selection, most do not carry out such robustness checks. OECD (2012), for instance, defines urban cores as high-density clusters of contiguous grid cells of 1 km^2 with a density of at least 1500 inhabitants per km^2, but use a lower density threshold of 1000 people per km^2 for Canada and the United States given that, in such countries "several metropolitan areas develop in a less compact manner" (OECD (2013) p. 3). These data processing choices make cross-country and regional inference non fully comparable.[12]

In response to the weaknesses of the first two approaches, recent work has moved towards a more harmonized definition of cities and rural areas based on the use of geo-referenced population datasets. Uchida and Nelson (2008), for instance, developed the agglomeration method to define a city, which relies on population size and density thresholds, as well as a travel time radius of a "sizable" settlement.[13] In turn, Dijkstra and Poelman (2014) follow a spatial approach and develop a cluster method defining two types of urban areas: *high density clusters*, with threshold values of 1500 inhabitants per km^2 and 50,000 people, and *urban clusters*, with thresholds of 300 inhabitants per km^2 and 5000 people.

This suggests that the geo-referenced data approach suffers from similar drawbacks as the other approaches, as it still relies on contiguity criteria as well as on thresholds that are subject to arbitrariness (see Roberts et al., 2016 for a discussion). Another pitfall that is common to all approaches is the perpetration of sources of measurement error along the construction process. For instance, the existing methods take into account national definitions of urban areas which are not comparable across countries and time, or geo-referenced population datasets that depend on censuses and therefore rely on each country's census minimum collection unit size, which in turn determines the data's spatial resolution.[14]

More recently, and to overcome the drawbacks of the traditional approaches to measure city size and urban area, researchers have started to use remote-sensing techniques. The next sub-section focuses on the use of the incidence of light at night to delineate the size and shape of cities.

2.1. Remote sensing approach and the use of nighttime light to measure urban areas

Satellite imagery and remote sensing techniques have been widely used to measure physical quantities related to urbanization and urban footprint. For instance, Burchfield et al. (2006) use satellite imagery of land cover and land use in U.S. from 1976 to 1992 to build boundaries of contiguous areas with similar land cover. They find that the extent of land use remained roughly unchanged during this period. More recently, Saiz (2010) uses satellite-generated data on terrain elevation and

[11] For a discussion on the topic see Duranton (2015).

[12] Another not so common approach to define metropolitan areas relies on non-economic criteria including concepts such as "sense of belonging", as captured through survey data.

[13] To that end, Uchida and Nelson (2008) used data from the Center for International Earth Science Information Network's (CIESIN), the Global Rural-Urban Mapping Project (GRUMP) and Landscan geo-referenced population datasets.

[14] One example is the Gridded Population of the World database (GPW), produced by the Socioeconomic Data and Applications Center (SEDAC) from NASA, and hosted by CIESIN. The data rely on an areal-weighting method (known as uniform distribution or proportional allocation method) to disaggregate population from census units into grid cells through the simple assumption that the population of a grid cell is an exclusive function of the land area within that pixel (Doxsey-Whitfield et al., 2015). The first disadvantage of using an areal-weighting disaggregation method is that "the precision and accuracy of a given pixel is a direct function of the size of the input areal unit. Consequently, for countries where the input units are quite large, the precision of population estimates for individual pixels within that unit can be degraded" (Lloyd et al., 2017). For example, the average input unit resolution for developed regions in the GPW data (version 4) is 944 km^2. Contrast that with an input unit resolution of 3518 (4,700) km^2 for middle-income (developing) regions (de Sherbinin and Adamo, 2015). Accuracy is thus proportional to development and introduces non-classical measurement error. To tackle this problem, CIESIN developed the Global Rural Urban Mapping Project (GRUMP) which, pretty much in the spirit of BEAM, assigns population over grid cells according to nighttime light data. However, GRUMP does not perform all the necessary corrections to the nighttime light data which we describe in this section, thus potentially generating biased estimates.

presence of water bodies to estimate the amount of developable land in U.S. metropolitan areas, and finds that inelastic housing supply is highly correlated with geographical constraints.

Perhaps the most important contribution to cross-country comparisons of urban footprint using satellite imagery is the *Atlas of Urban Expansion* (AUE), produced by the Marron Institute of Urban Management at New York University, in collaboration with UN-Habitat and the Lincoln Institute of Land Policy. AUE uses high resolution satellite imagery together with disaggregated census data to estimate, for 200 cities in the world, urban footprint, urban built-up area, population density and several indicators that characterize the specificities of the urban outlook of cities. These include the shares of urban infill, leapfrog, and extension for the years 1990, 2000 and 2014 (Angel et al., 2012).

While valuable and highly precise, the approach followed by the AUE is arguably very costly. There are two reasons for this. On the one hand, AUE uses as input various sources of information, including high resolution satellite images from Landsat and Google Earth, disaggregated census data, and specific surveys obtained from local researchers measuring different land uses, property regimes, affordability of housing and attributes of properties available for sale or rent. On the other, producing the detailed portfolio of AUE indicators is computationally demanding. This may partly explain AUE's reduced sample of cities. Furthermore, in the same fashion of aforementioned approaches, census and land use data cannot be pooled to provide suitable comparable observations due to national differences on census methods, accuracy and capture times.

One alternative is to use the Nighttime Light longitudinal data available from the Defense Meteorological Satellite Program (DMSP) of the National Center for Environmental Information (NOAA). DMSP satellites take cloud-free annual composites images to produce digital remotely sensed images of the world's light at night, that unlike a photograph, can be manipulated to extract information.[15] Each image corresponds to a 30 arc-second pixel (roughly 1 km^2 at the equator) containing a specific *Digital Number* (DN). The DN is assigned according to the pixel luminosity, and varies between 0 to 63.

Among other uses, nighttime luminosity can help predicting urban settlements.[16] This has been done at least since Elvidge et al. (1997). Indeed, night lights have been shown to correlate accurately with population density in the U.S. (Sutton et al., 1997), to identify urbanization rates (Imhoff et al., 1997), to estimate global population counts (Sutton et al., 1999), and even to single out low density settlements (Elvidge, 2000). Moreover, the use of night-time satellite imagery to derive urban boundaries has been validated both for developed and developing countries (Henderson et al., 2003), and several studies have used this tool to study urbanization changes at different geographical scales (see He et al. (2006), Zhang and Seto (2011) and Liu et al. (2012), among others). Harari (2017) uses night time lights to classify the shape of Indian cities and finds that city compactness is correlated with lower wages and higher housing rents.[17]

Nighttime luminosity can be combined with high resolution Landsat images for more accurate predictions of urban construction and population counts. For instance, Landsat images can be used to extract green areas and bodies of water to refine the measures derived from night lights (see for instance the work of Goldblatt Ran et al. (2018) for India, Mexico and the U.S.). However, nighttime light data suffers from a series of problems that, in the absence of the appropriate corrections, affect its reliability and comparability across space and time (Abrahams et al., 2018; Elvidge et al., 2009; Liu et al., 2012; Zhang and Seto, 2011).

For instance, factors such as the exact position of satellites at the time of capturing the information, and the satellite-specific degradation of the capturing sensors imply that recorded images can suffer from misalignment, and cause DN values within a satellite to vary across time, and across satellites within the same year for reasons other than the actual luminosity intensity of pixels. Specifically, DMSP has five satellites that have captured nighttime light images over the period 1992 to 2013. For several years two satellites capture annual light composites. However, first abnormal fluctuations appear in DN values for different years derived from the same satellite. Second, the number of lit pixels differs between two satellites for the same year. Third, the number of lit pixels derived from the same satellite decreased abnormally between different years. In addition, nighttime light suffers from blurring, which means that light "overglows" and cities seem to be bigger than they actually are, especially if they are more luminous (because light spatial correlation increases with the degree of luminosity). Finally, this data suffers from geographic misalignment within and between satellites because the images can vary depending on the exact position that the satellite had when it captured the information.

The literature has proposed several ways to address each of these issues, albeit in a decentralized fashion. An important contribution of this paper is to propose general and replicable procedure to apply the known correction techniques while relying on minimal cleaning and processing assumptions to estimate the urban area of cities in the entire world. The next subsection describes the data sources as well as the cleaning and processing procedures.

2.2. Data

Following the conceptual definition of cities as the spatial integration of social and economic activity, as well as the growing literature on how night lights are strongly associated with economic performance and development, this paper uses 1 km^2 images of luminosity at night to compute the size of metropolitan areas for the entire world, in 2000 and 2010. To this end, we rely on average visible nighttime light raster files and frequency images provided by NOAA.[18] Our methodology can be divided in four steps: i) identification of urban cores with more than 50,000 inhabitants; ii) the cleaning and processing of nighttime light data; iii) the classification of pixels to define the existence of a city and its extension; and iv) the estimation of population zonal statistics. In the rest of this section we provide a detailed description of each such step.

2.3. Methodology

2.3.1. Identification of urban cores

An important challenge to estimating the size of cities greater than 50,000 inhabitants is identifying where these cities are located to begin with. Indeed, there is no global dataset that tracks over time the creation of urban cores around which metropolitan areas may form. To address this concern we take two approaches. First, following Roberts et al. (2016) and Dijkstra and Poelman (2014) we use a cluster method that follows a spatial/demographic algorithm to yield a consistent and comparable cross-country measure of the presence of urban areas. The resulting set of urban cores are used as reference points from which

[15] Satellite images take into account only persistent lighting, thus discarding ephemeral events including fires.

[16] Luminosity satellite data has been used as a proxy for economic activity across countries, at the sub-national level and for very small areas (Bleakley and Lin, 2012; Henderson et al., 2011; Lowe, 2014; Martinez, 2018; Michalopoulos and Papaioannou, 2013; Storeygard, 2016; Weidmann and Schutte, 2016). The literature on night lights as a proxy for development is summarized by Donaldson and Storeygard (2016). These data have also been used, in conjunction with high-resolution daytime satellite images and machine learning tools, to predict sub-national poverty (Jean et al., 2016).

[17] For a thorough discussion on the use of nighttime light data to evaluate urban coverage and change see Goldblatt Ran et al. (2018).

[18] The raster images of interest are those from the F14-2000, F15-2000, F18-2010 satellites-years. We do not rely on radiance-calibrated images given that, since we do not perform any within-city inference, we are not concerned about top-coding (the fact that high light values end up at the top of the 0–63 scale), but only about the emission of light above certain frequency threshold. Hsu et al. (2015) discuss how radiance-calibrated nighttime lights deal with top coding in high luminosity areas, such as city centers.

BEAM's methodology is applied to detect urban areas of at least 50,000 people.

Specifically, to identify the approximate location of urban cores, we classify population grid cells according to two thresholds: a population density threshold of 300 people per km^2, and a total population count threshold of 50,000 inhabitants.[19] To that end we use the 2015 Global Human Settlement (GHS) geo-referenced population raster files, that spatially disaggregate population across the surface of the world.

Next, we construct core centroids, from which we generate buffers of 130 km radius.[20] The resulting area constitutes the space in which we then carry out the correction procedure described in Section 2.3.2, and particularly deblurring. Indeed, Abrahams et al. (2018) suggest that the deblurring procedure described below needs to be limited to a particular geographic area and, due to top censoring, additional area needs to be added since it is typically less precise when it approximates the edges of the space considered.

The second approach to identify cities is using as potential urban cores the universe of cities identified by the AUE. As noted by Angel et al. (2005), this list is not necessarily complete, and thus we use it to complement the population cluster method described above.[21,22]

2.3.2. Cleaning and processing nighttime light

DMSP nighttime light data suffers from blurring. Light spreads far beyond urban built-up coverage, and brighter pixels generate larger blurs. Water bodies also tend to spread light across space. We rely on Abrahams et al. (2018)'s two-step deblurring filter. The filter specifies a pixel stability threshold ranging from 0 to 100, where all pixels with a frequency lower than the stated threshold are set to zero (that is, no nighttime light is assumed there). A 0% threshold implies that a pixel that appeared lit in just one satellite picture during a specific year is coded as a lit pixel. A 100% threshold implies that to be coded as lit the pixel needs to have light in every satellite picture within a year. We do not hold a strong prior in favor of a specific pixel stability threshold, and thus estimate BEAM based on various thresholds (0%, 20% and 50%).[23] It is worth noting that deblurring corrects for spatial correlation in the measurement of city size.[24]

After deblurring the night light images, we apply a number of additional correction procedures to ensure a reliable estimation of city size. First, because images recorded by the same satellite may not be comparable across years, we apply Wu et al. (2013)'s inter-calibration procedure.[25] Second, as noted by Zhao et al. (2015), nighttime light rasters obtained from different satellites (as well as for different years from the same satellite) tend to be shifted a couple of pixels, reducing comparability across sources and over time. Following these authors, we then fix a reference image (specifically that of satellite/year F142001) and shift all the other nighttime lights rasters to be aligned with it.[26] Third, comparing the information recorded by two different satellites (F14 and F15) within a given year allows us to remove any intra-annual unstable lit pixels, that is those detected as having a positive DN by only one satellite. In addition, we assign to every stable lit pixel the average DN value of the two satellites, hence producing an intra-annual stable *composite* for each sample year.[27]

2.3.3. Pixel classification and separation of very large metropolitan areas

After the pre-processing described in the previous section, we classify pixels into urban polygons according to the following 3-stage process. First, we identify pixels with a DN > 0 (i.e. any lit pixel, based on the pixel stability threshold of 0%, 20% or 50%) and drop those with $DN = 0$. Second, we classify as a city any conglomerate of at least 1 km^2 of contiguous stable lit pixels.[28] Third, we merge the country shapefiles of Natural Earth and DIVA-GIS to assign cities to existing countries.

In several cases, estimated metropolitan areas are big enough to encompass two or more existing cities that are connected by somewhat narrow corridors of urbanized areas (as identified by BEAM, but beyond the luminosity of connecting highways). This is the case, for instance, of the metropolitan Northeast corridor of the U.S. from New York City to Washington D.C., the Chicago metropolitan area and its continuity up to Milwaukee, the Buenos Aires Metropolitan Area and its extension to La Plata or Mexico City's tie to the nearby city of Toluca.

Our approach contrasts to that of other datasets such as AUE, which rely on administrative boundaries to split cities. To an extent, such splits contradict the very purpose of BEAM, namely identifying and measuring the extent of metropolitan areas as clusters of economic activity that become apparent in nighttime luminosity. However, it can be argued that splitting large cities that are commonly perceived as distinct favors comparability of BEAM with other datasets. Clearly, the challenge is to come out with a procedure that makes such splits automatic, without handpicking neither cities nor specific splitting areas. We thus split light-connected metropolitan areas if the width of the connecting urban corridor identified in our procedure is less than a certain threshold. For robustness we set such a threshold from 1 to 5 km in increments of 1 km.[29] Lastly, we clip oceans and inland water bodies using Natural Earth v2.0.0 data with a 1:10 m scale.

This three step procedure -pixel identification, urban classification and clipping-, and the combination of the five splitting thresholds (1 to 5 km) and the three pixel stability thresholds (0%, 20% and 50%), yields 15 different versions of BEAM, both for 2000 and for 2010. In the rest of the paper, we show the robustness of our results to these variations. Ultimately, based on the paper's replication material, each researcher

[19] One km^2 is equivalent to 0.009 x 0.009 decimal degrees.

[20] According to our estimates, the metropolitan area with the largest distance from its centroid to its farthest perimeter point is Los Angeles, with 110 km. Thus, a 130 km radius seems conservative.

[21] Fig. A.1 and A.2 in the Appendix plot the centroid of the urban cores identified using the AUE and GHS respectively, and compare them to the centroid of the metropolitan areas found by BEAM after applying all the corrections described in Sections 2.3.2 to 2.3.4. Clearly, some of the cores listed by these two sources are not ratified by BEAM. These are mainly located in Sub-Saharan Africa and China (Fig. A.1 and A.2), and the Amazon basin, South Asia and the Middle East (Fig. A.2). This suggests that most of these agglomerations are more rural than required by BEAM. We rule out alternative explanations, such as low energy consumption (see Table A.2) or high incidence of clouds that affect the light-detecting capacity of the satellites (see Fig. A.3).

[22] A third approach would identify the areas that appeared lit using the non-deblurred nighttime light images (after applying all the corrections other than deblurring explained in Sections 2.3.2 to 2.3.4). Our estimates are robust to using this third measure, which we used for an early version of the data. We, however, do not follow this approach in the current version of BEAM as it turns out that the approaches described in the text are more precise.

[23] Abrahams et al. (2018) note that the 20% threshold removes transitory lights without affecting city size for the case of 15 cities.

[24] A different approach to correct the blurring problem when measuring cities with night lights is to use luminosity thresholds in an attempt to screen rural areas out. However, as discussed, thresholds are arbitrary and data are not comparable when different thresholds are used for different regions, as in the case of Liu et al. (2012).

[25] Inter-calibration parameters derived by Wu et al. (2013) are based on blurred night light raster files. Future work needs to estimate new parameters using global deblurred images.

[26] As a result, satellite F142000 is unchanged, F152000 is moved left 1 pixel and up 1 pixel, and F182010 up 1 pixel.

[27] Liu et al. (2012), using the same procedure for inter-annual corrections, assume that cities cannot *reduce* their size overtime, and thus do not remove pixels unidentified by one of two satellites. While this is probably true in most cases, we do not make any such assumption a priori.

[28] We use the rook contiguity criterion, that is when polygons share a common side.

[29] Specifically, we first identify all the vertices of the perimetral pixels that compose a city's polygon and draw straight lines connecting all non-contiguous pairs of vertices. We then identify city areas in which lines are shorter than the threshold, and "turn off" the entire clustered area, thus splitting the original city polygon into two. This splitting procedure does not depend on population counts or population density, but only on a city's shape.

can estimate their own dataset, using the thresholds that she considers adequate (including thresholds other than those used to generate the data for this paper).

The resulting city geo-referenced polygons can be combined with existing population datasets to compute various city-level statistics of interest. Specifically, in the next subsection we discuss how BEAM can be used to compute urban population counts and population density.

2.3.4. Population zonal statistics

Several population raster files exist to estimate population counts for a given city and over time.[30] For our baseline results, we generate zonal statistics (city-level population counts and population density) using the Landscan geo-referenced population gridded raster files (for the years 2000 and 2010). However, for robustness we also use GHS (for 2000) and WorldPop (for 2000 and 2010). Unfortunately, the latter only covers Latin America, Africa and Asia, while the former does not have data for 2010.[31]

Overlaying BEAM to population rasters allows us to limit the sample of cities to those with at least 50,000 inhabitants. We do so for two substantive reasons. First, this ensures we exclude other light emission zones such as oil refineries. Second, we are interested in revisiting stylized facts about urbanization and the size distribution of cities (including Zipf's law and Gibrat's law) that have been recently studied using cities defined with similar minimum population thresholds.[32] Available datasets that define cities or metropolitan areas using the same population threshold include the U.S. MSAs and the datasets created by Dijkstra and Poelman (2014) for Europe and by Roberts et al. (2016) for the entire world.

2.4. Benefits of BEAM

Our measure of the size or metropolitan areas based on nighttime lights has several benefits in terms of accuracy and computational practicality. First, it allows a comparable analysis of stylized facts about urbanization across different regions of the world, with potentially non-homogeneous administrative definitions of cities. Second, it does not rely on assumptions about people's mobility, nor on the availability of such data (like the measures based on commuting patters). Third, the methodology does not assume an arbitrary population density threshold, or a threshold expressed in terms of time and number of travels. We do make use of pixel stability and metropolitan areas split thresholds, but show the robustness of our results to thresholds that vary across a range of different values. Fourth, while the computational demand of BEAM is still high, it differs substantially from that of other methods, including the use of construction coverage. Fifth, our methodology does not establish a population threshold ex-ante. Rather, the population count of an area is estimated *after* the creation of the city's urban extension. This allows researchers to vary the minimum city size depending on their specific interests and also guards BEAM from potential political manipulation.[33]

Lastly, and most importantly, luminosity has been extensively proven to be a reliable proxy of economic activity, and so it allows us to identify clearly integrated urban markets. As noted for the case of Bogota by Duranton (2015), the city is not officially constituted as a metropolitan area even though there is no discontinuity with several nearby municipalities, including large towns such as Soacha (in the south) and Chia (in the north). The lack of administrative integration does not, however, affect the economically relevant estimates of BEAM.

It should also be noted that the cleaning and pixel processing algorithm is transparent and simple, and it is flexible enough to allow for changes that any researcher may consider relevant.[34] According to Duranton (2015), transparency, clarity and flexibility are desirable characteristics of these type of measures.

Privileging computational practicality and comparability across regions despite an arguably small cost in terms of accuracy (as DMSP nighttime lights have a resolution to the 1 km^2), we chose not to complicate our measure by, for instance, subtracting green areas and non-built up areas, unless they do not form part of the conglomerate of lit pixels.

3. Basic city-level statistics

Based on the 15 BEAM datasets that we generated for the purpose of this paper, and using Landscan spatial population estimates, Table 1 presents the minimum and maximum number of resulting metropolitan areas of at least 50,000 in both 2000 and 2010. It does so for each one of the seven regions defined by the World Bank as well as for the entire world.[35]

How does the estimated number of cities compare to the equivalent figure of existing datasets? We now discuss some basic comparisons and refer to Appendix A.2 for further details.

3.1. Comparison with existing datasets

We start by comparing the number of cities identified by BEAM to the urban areas listed by the AUE (Angel et al., 2005). For example, using the version of BEAM computed for 2010 with a deblurring threshold of 0% and a metropolitan area-split threshold of 5km, we find that BEAM accounts for 90% of the cities identified by the AUE. BEAM misses urban areas that are primarily located in Sub-Saharan Africa and China. In the former, cities tend to be more rural than what is required by BEAM (see Fig. A.1).

This is also the result of the comparison of BEAM to the list that can be estimated using a population cluster method and the GHS population raster file. As illustrated in Fig. A.2, BEAM identifies less metropolitan areas relative to the urban cores listed by GHS in Sub-Saharan Africa, South Asia and the Amazon basin. Alternative explanations of this mismatch include the potential low energy consumption (and thus low nighttime light activity) of these relatively poor areas, or the presence of cloud clusters that affect the light-detecting capacity of the satellites. As discussed in detail in Section 5, neither of these seems to be the case. Specifically, as we show in Table A.2 in the Appendix, there is no correlation between the average size of all the BEAM-estimated metropolitan areas of a country and its per capita energy consumption, controlling for country and time fixed effects. Moreover, Fig. A.3 in the Appendix suggests that there is no high cloud presence in regions were the mismatch is larger.[36]

[30] Section A.3 of the online appendix provides a description and comparison of the main geo-referenced population datasets and discusses their specific advantages and disadvantages.

[31] Other population datasets such as the Gridded Population of the World (GPW) and the Global Rural Urban Mapping Project (GRUMP) are not used for the reasons described in the appendix but researchers who download our replication files could easily use them.

[32] In particular, the Zipf's law is more likely to hold in the upper tail of the city-size distribution. In the left-tail, it is much more sensible to the definition of what constitutes a city (Ioannides and Skouras, 2013).

[33] This is a concern that arises when utilizing pre-defined urban cores, as noted by Duranton (2015).

[34] Indeed, for precisely that purpose, we provide detailed replication material accompanying this paper.

[35] The minimum and the maximum number of cities in each region/year reported in Table 1 results from a different combination of pixel stability and metropolitan area-split thresholds, and thus we encourage readers to avoid computing city count growth rates based on this Table. Instead, Table A.1 of the Appendix reports the number of metropolitan areas across the 15 versions of the data.

[36] Additionally, recall that BEAM relies on light contiguity and population counts within those boundaries. Instead, methodologies that are solely based on population rasters to estimate the size of cities usually aggregate otherwise spatially scattered population counts around census minimum collection units.

For the case of the U.S. we compare the number of metropolitan areas estimated by BEAM to this country's Metropolitan Statistical Areas (MSAs) identified for 2010. BEAM identified cities in 373 of the 389 MSAs, and the residual corresponds to MSAs that do not have more than 50,000 inhabitants within a continuous urban space (see Fig. A.4 in the Appendix).[37] BEAM also identifies *all* the cities listed by AUE for the U.S. in 2010 (see Fig. A.1).

Further, in Table A.3 of the Appendix, we compare the number of people living in metropolitan areas of different sizes in the U.S., contrasting the MSAs to BEAM. We find that the distribution of people across city sizes is remarkably similar in the two datasets.

3.2. Cross regional statistics

We now report basic stylized facts about urbanization. This is made in relation to three dimensions: the spatial extent of cities or metropolitan areas, urban population and population density, which is the ratio of population to area. These are summarized in Fig. 1.[38] Panel A shows the distribution of the log of the size of metropolitan areas in each region, as estimated by BEAM.[39] Regions are organized in an ascendant manner according to the median size of metropolitan areas in 2000.

South Asia and East Asia and the Pacific were the regions with the lowest median city size in 2000. Europe and Central Asia, Sub-Saharan Africa, Middle East and North Africa and Latin America and the Caribbean had somewhat larger cities, and North America had significantly larger cities compared to any region.[40] By 2010, all regions grew in area.

As described in Section 2.3.4, we overlay BEAM with the Landscan geo-referenced population datasets for the years 2000 and 2010. This allows us to compute urban population levels for the estimated metropolitan areas in every region. Interestingly, Panel B of Fig. 1 shows that the distributions of urban population were remarkably similar across regions in 2000. In 2010 North America and Europe had a slightly lower population medians than the rest of the world, and South Asia had a slightly higher median.

Panel C, in turn, reports population density, which by and large are the mirror image of urban areas, precisely because of the similarities in urban population levels across regions: both in 2000 and 2010 South Asia, East Asia and Sub-Saharan Africa, were the regions with the highest median population density. North America was, to a large extent, the region with the lowest urban density.

The evidence presented in panel C of Fig. 1 is consistent with the existence of a negative association between population density and degree

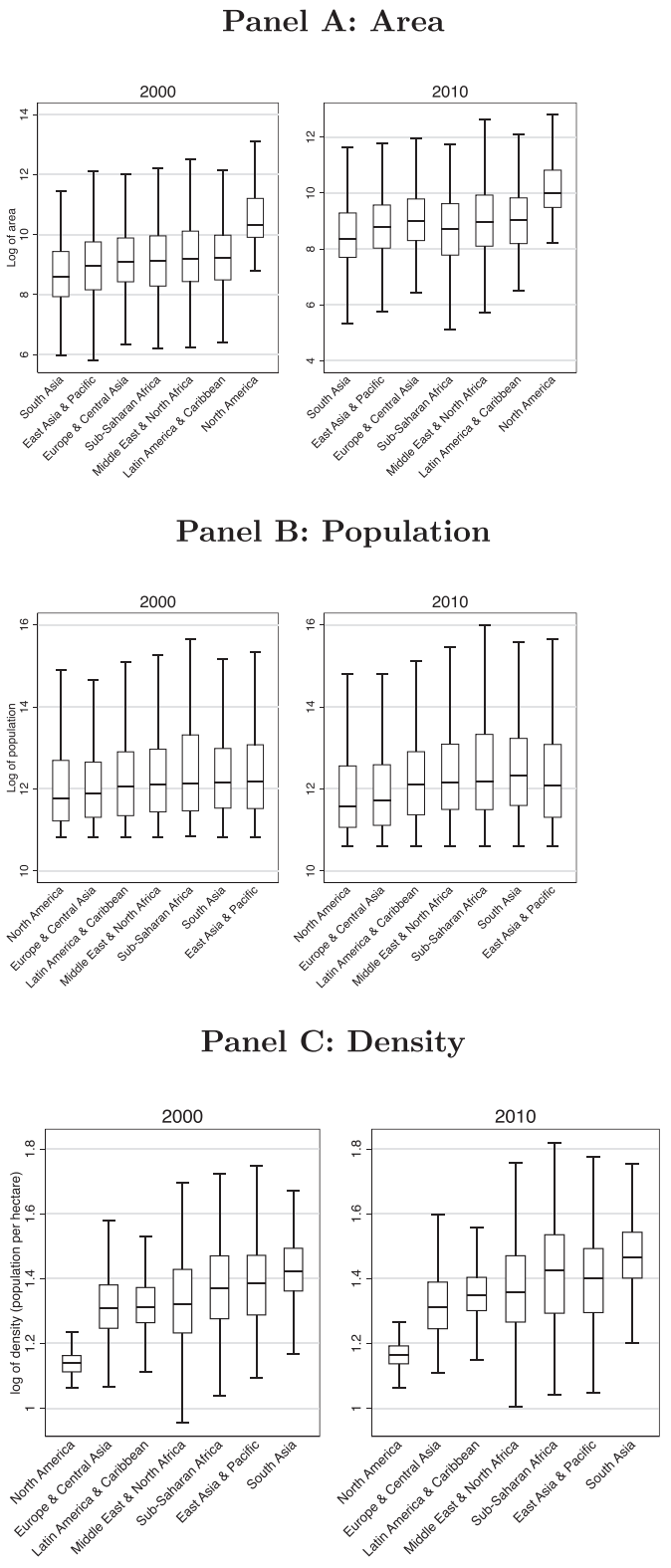

Fig. 1. City size distribution by region. **Notes:** The box plot quantities are explained in footnote 39. In Panel B, population is estimated using *BEAM* and Landscan georeferenced population dataset, with pixel stability of 0% and splitting cities with connecting corridors less than 5 km wide. In Panel C, the estimated median density -in levels- per region in 2000 [2010] is: in North America, 4 inhabitants per hectare [5]; in Europe and Central Asia, 17 [18]; in Latin America and the Caribbean, 20 [24]; in Middle East and North Africa, 20 [226]; in Sub-Saharan Africa, 28 [53]; in East Asia and Pacific, 33 [38]; and in South Asia 41 [59].

This is likely to identify cities that would otherwise not be identifiable using light, given population thresholds of at least 50,000 inhabitants.

[37] For example, consider the MSA of Farmington in the Northwest corner of the state of New Mexico. This MSA is composed by various non-contiguous cities, including Aztec (6,763 inhabitants in 2010), Bloomeld (8,112) and Farmington (45,877). While the sum of people across these cities is greater than 50,000 (and indeed adds up to 130,044 in 2010 if all the towns are added), no individual urban core exceeds this number.

[38] This figure was made using a pixel stability threshold of 0% and a metropolitan area split threshold of less than 5 km. Hereafter and for comparison purposes, all figures use the same thresholds. However, to a large extent ordinal results hold irrespective of the threshold selection.

[39] The quantities of interest included in the box plot are: the median of the distribution (horizontal line inside the box); the 25^{th} and 75^{th} percentiles (box limits); upper and lower adjacent value (box whiskers, defined as the box upper (lower) limit + (-) 3/2 of the 75^{th} to 25^{th} percentiles gap). To facilitate the visual inspection of the distributions, we remove outliers (values outside the whiskers).

[40] Hereafter and to save space without scarifying clarity by using the World Bank region acronyms, we will write "East Asia" to refer to East Asia and the Pacific; "Europe" to refer to Europe and Central Asia; "Middle East" to refer to Middle East and North Africa; and "Latin America" to refer to Latin America and the Caribbean.

Table 2
Urbanization rate by region.

	2000		2010	
	Min (1)	Max (2)	Min (3)	Max (4)
East Asia and Pacific	0.42	0.46	0.41	0.48
	(0.32)	(0.33)	(0.34)	(0.34)
Europe and Central Asia	0.42	0.53	0.35	0.50
	(0.18)	(0.18)	(0.16)	(0.19)
Latin America and Caribbean	0.49	0.56	0.52	0.60
	(0.21)	(0.21)	(0.52)	(0.20)
Middle East and North Africa	0.58	0.65	0.57	0.72
	(0.29)	(0.27)	(0.31)	(0.22)
North America	0.60	0.63	0.68	0.72
	(0.12)	(0.13)	(0.04)	(0.72)
South Asia	0.16	0.21	0.15	0.23
	(0.12)	(0.14)	(0.10)	(0.13)
Sub-Saharan Africa	0.14	0.15	0.14	0.19
	(0.15)	(0.16)	(0.14)	(0.15)

Notes. Columns report minimum and maximum urbanization rates (and standard deviation in parentheses). Figures are computed using *BEAM* and Landscan georeferenced population dataset, considering cites with more than 50,000 inhabitants, applying pixel stability thresholds of 0%, 20% or 50% and city split thresholds of 1 km, 2 km, 3 km, 4 km, or 5 km.

of development. This can have two complementary explanations. On the one hand, increases in the average income levels are associated with declines in urban density. This is explained by a positive relationship between income and the demand for residential space. On the other hand, the relatively high levels of population density in developing regions may be a reflection of the recent, accelerated and informal urbanization process, experienced by these regions since the mid-twentieth century. In fact, much of the high density of developing cities is explained by the incidence of informal settlements (Jedwab and Vollrath, 2019). In the next subsection, we discuss further evidence on the relationship between urban rates and income.

3.3. Urbanization rates and economic development

Combined with geo-referenced population data, BEAM allows us to compute urbanization rates. As mentioned, in this paper we do so using Landsat population rasters. Note that these are computed conditional on cites having at least 50,000 inhabitants. We do so for the 15 versions of BEAM produced in this paper, for the seven world regions defined by the World Bank and for both 2000 and 2010. We report the resulting figures (along with their standard deviation) on Table A.4 in the Appendix, and summarize the findings on Table 2 by reporting the minimum and maximum estimated urbanization rates for each region/year. The dispersion between the minimum and the maximum is reassuringly rather small (less that 10 percentage points) in five out of the 7 regions: East Asia, Latin America, North America, South Asia and Sub-Saharan Africa.[41] Europe and Middle East have somewhat larger dispersion (up to 15 percentage points) and thus, in these regions, urbanization rates computed using BEAM are less robust to changes in the pixel stability and the city-split thresholds.

Urbanization levels computed with BEAM are smaller than the figures that other organizations compute using cities administrative levels. This difference is in part mechanic, as BEAM focuses on cities of a minimum size. North America has the highest urbanization levels (around 60–70 percent), and also the largest growth rate between 2000 and 2010, irrespective on whether the minimum or the maximum figure is used to compute the growth rate (around 8–9 percentage points). Middle East and Latin America are the next two regions with the highest urbanization rate (and also positive inter-period growth rate, albeit smaller than that of North America). South Asia and Sub-Saharan Africa are the least urbanized regions (around 15–20%) and their urban population growth rate is very small. Interestingly, while Europe has relatively high urbanization levels (around 40–50%), urbanization seems to have slightly decreased from 2000 to 2010.

We now study the correlation between a country's urbanization rate and its (log) GDP per capita. We do so separately for 2000 and 2010, controlling for region fixed effects. We also study this relationship exploiting the two year panel, to be able to control for country and year fixed effects.[42] We estimate Eq. (1) for the cross sections and Eq. (2) for the panel.

$$\text{Urb.Rate}_i = \alpha + \beta \log(\text{GDPpc})_i + \delta_R + \varepsilon_i \quad (1)$$

where δ_R are region fixed effects and the error term is clustered at the region level.

$$\text{Urb.Rate}_{it} = \alpha_i + \gamma_t + \beta \log(\text{GDPpc})_{it} + \mu_{it} \quad (2)$$

where α_i and γ_t are respectively country and year fixed effects, and the error term is clustered at the country level.

We report the estimated coefficients on Table 3 in sets of three columns for each version of BEAM. The first (second) column looks at the 2000 (2010) cross section, and the third at the two-year panel. The relationship between the two variables is positive and significant in each one of the two cross sections, and the magnitude is fairly similar in both years and across the version of BEAM used to estimate the regression model. Specifically, in 2000 (2010) a 10 percent increase in the log of GDP per capita is associated with urbanization rates that are between 1.3 (1.4) and 1.5 (1.7) percentage points higher. The top panel of Fig. 2 presents this relationship graphically.

Interestingly, the relationship between GDP and urbanization rate turns negative once country-specific time-invariant factors, as well as country level shocks are accounted for by the two-way fixed effects (columns 3 and its multiples of Table 3). The magnitude is, however, rather small: a 10 percent increase in the log of GDP per capita is associated with urbanization rates that are between 0.09 and 0.54 percentage points *lower* when the panel structure of the data is exploited. It is worth noting that this result is less precisely estimated, as only 3 out of the 15 point estimates are statistically significant in the presence of the clustered standard errors. In any case, this suggests that the apparently well-established stylized fact reported in the introduction of the paper –that urbanization is one of the most salient manifestations of economic development– needs to be interpreted cautiously, at least in the short run. According to our estimates, the relationship is, at best, statistically nonexistent.

These results hide interesting regional variation, which we examine by estimating augmented versions of Eqs. 1 and 2, that add an interaction term between log(GDP per capita) and regional indicators.[43] The results are reported on Appendix Table A.5. In the cross sections, we find that urbanization reacts stronger to GDP changes (the interaction term is positive) in East Asia, Latin America and North America. In the other regions the interaction term is not significant.[44] Perhaps more interestingly, in the two-year panel, we find that the effect of income of urbanization is a well measured zero (the interaction term is positive and of

[41] The dispersion is the lowest in both North America and Sub-Saharan Africa, suggesting that the robustness of BEAM is not correlated with a region's level of development.

[42] Alternatively we could estimate a first-differences specification, which yields the same point estimates but somewhat smaller standard errors.

[43] Each region dummy is interacted with income on a separate regression, in order to interpret the estimated coefficient of the interaction as the differential effect of income on urbanization in that specific region relative to the rest of the world.

[44] The described regional variation is portrayed graphically in Fig. 2, which plots for each cross section the correlation between GDP per capita and urbanization rate for the entire world as well as for a different region in each panel. The correlation coefficients for each region/year are reported in the table notes.

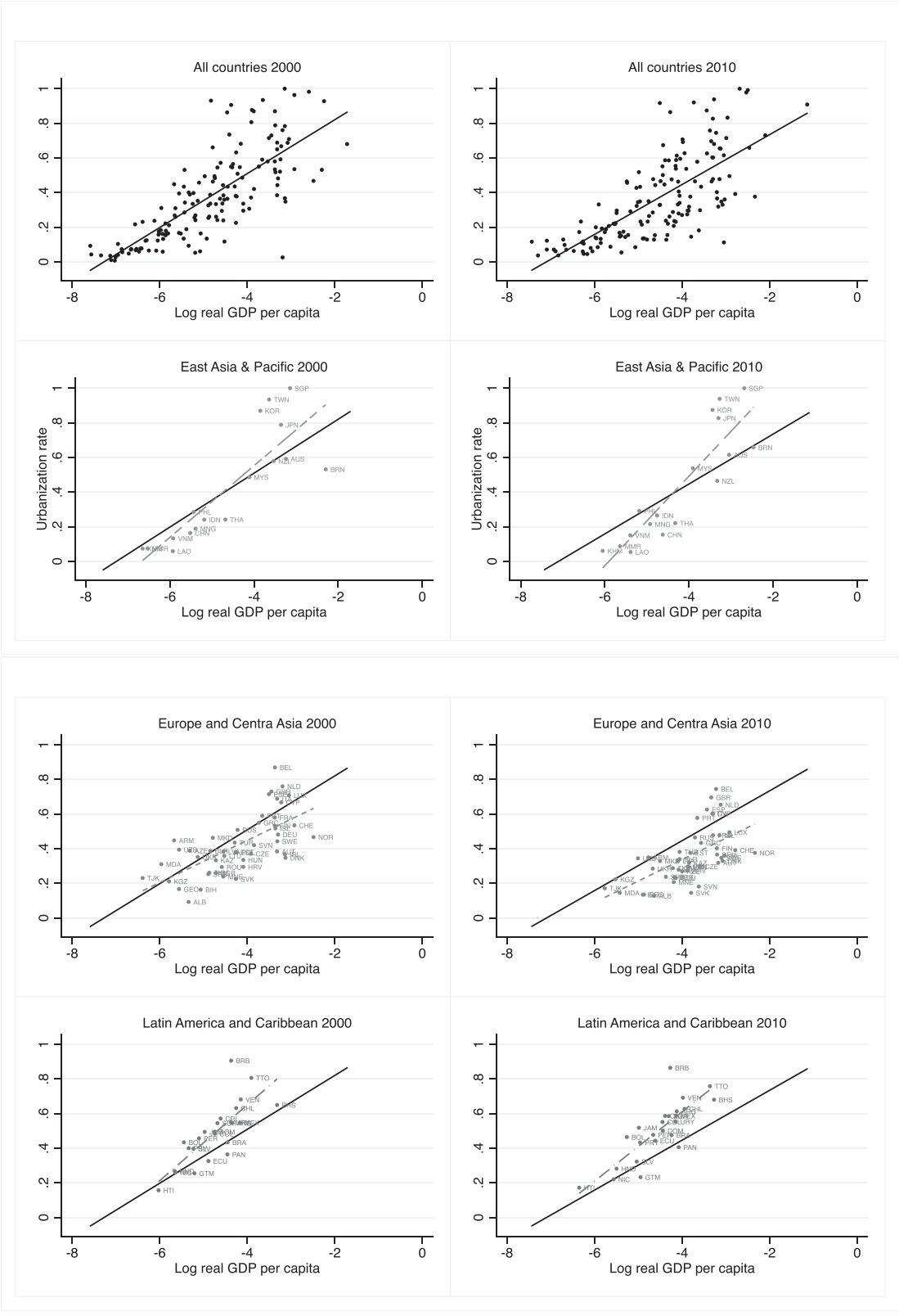

Fig. 2. Urbanization rate and GDP per capita. **Notes:** Each dot represents a country. The three letters next to each point are the ISO 3 country codes. Each line displays the regression of the logarithm of real GDP per capita on urbanization rate for either all countries or each of World Bank's regions. In 2000 [2010], the beta estimator (the solid line) for all the countries is 0.156*** [0.144***]; for East Asia 0.207*** [0.259***]; for Europe 0.122*** [0.124***]; for Latin America 0.219*** [0.200***]; for Middle East 0.132** [0.141***]; for North America 0.963 [0.399]; for South Asia 0.175** [0.133***]; and for Sub-Saharan Africa 0.102*** [0.092***]. The point estimates do not coincide with those in Table 3 since we do not include region fixed effects for the estimates of this figure. Significance-levels *** 1%, ** 5%, and * 10%, refer to two-sided t-tests. Real GDP is at constant 2011 national prices (source: Feenstra et al., 2015). Urbanization rates are estimated using *BEAM* and Landscan georeferenced population dataset, applying a pixel stability and cities' split thresholds of 0% and less than 5 km, respectively.

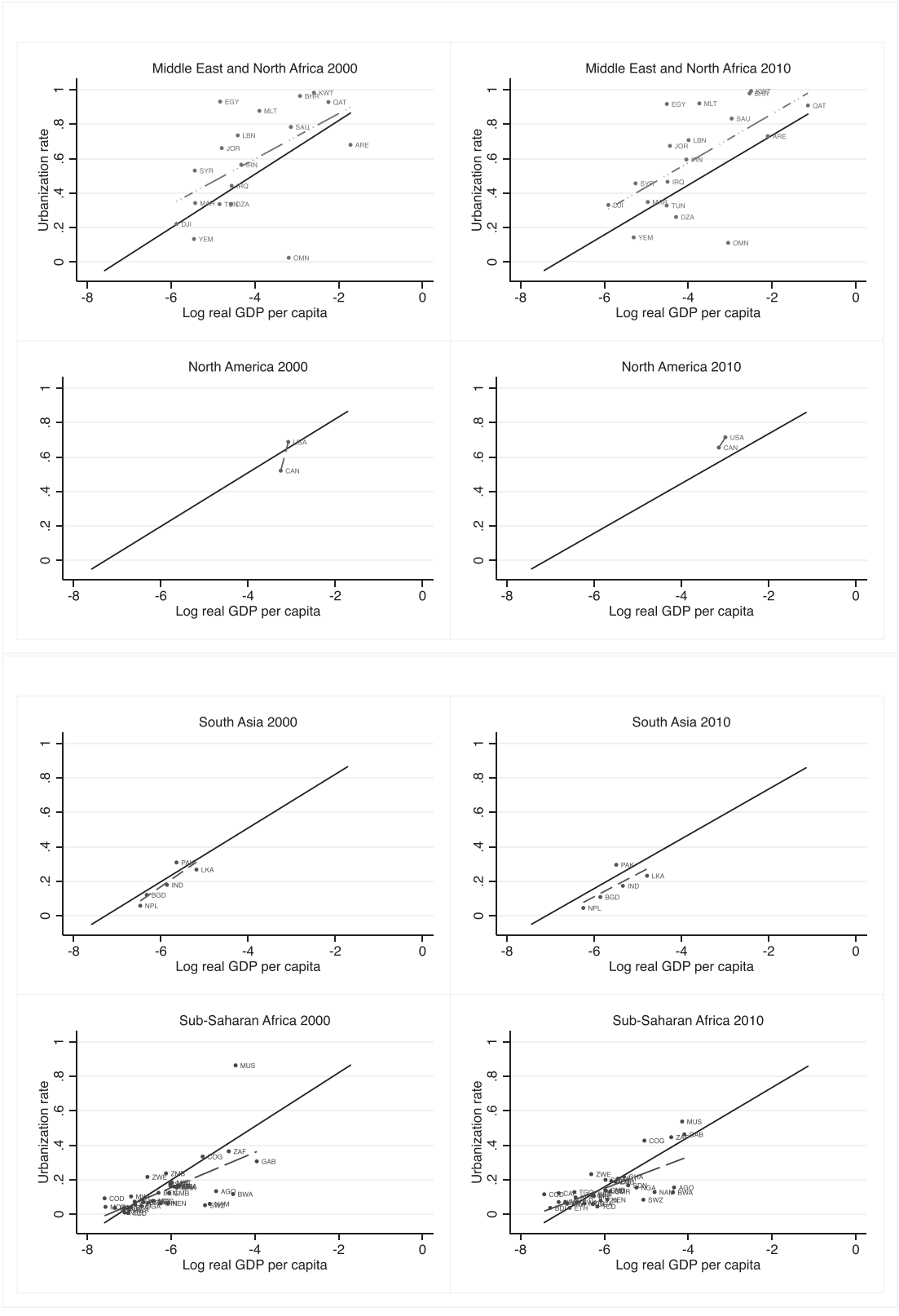

Fig. 2. Continued

the same magnitude as the negative non-interacted effect) in East Asia, Middle East and Sub-Saharan Africa; negative in Europe (the interaction term is negative and much larger in magnitude than the -positive- non-interacted term); and somewhat positive in Latin America and North America (in most specifications the interaction term is positive and larger in magnitude than the negative non-interacted term). This suggests that the negative and small effect reported on Table 3 when exploiting the panel structure of BEAM is likely driven by Europe, which

is by and large the region with the greatest number of countries in our sample of metropolitan areas (29%).[45]

4. City size distribution and growth

In this section we study the size distribution of cities, as well as the relationship between city size and growth rate for all the countries of the world based on a homogeneous and comparable measure. We do so by studying the extent of which BEAM-identified metropolitan areas can be described by Zipf's law and Gibrat's law, respectively. In essence, Zipf's law states that, if we consider the N cities from a country and sort them according to their size from 1 to N, the city in position *i* will have a population equal to that of the largest city divided by *i*. So, the second city will have half the population of the first, the third, one third, and so on (Gabaix, 1999a).[46] Zipf's law is also called the rank-size rule (Rosen and Resnick, 1980). Compliance with Zipf's law implies that cities' growth process is independent of their own size and is due to exogenous productivity shocks (Duranton and Puga, 2014). If this is true, then Gibrat's law -which states that the growth of urban areas is independent from their size- should also hold.[47]

While the size distribution of cities is one of the most studied phenomenon by urban economists, comparisons across countries or regions are limited by the lack of homogeneous data (Duranton and Puga, 2014). For instance, in a recent paper, Chauvin et al. (2017) explore whether the known facts about urbanization in the U.S. also hold in developing countries, specifically Brazil, China and India. To this end, the authors include tests of Zipf's law and Gibrat's law, comparing the four countries using different definitions of city's extension, a limitation the authors are aware of and upfront about when recognizing that the definitions used are "debatable", although probably the best available at the time of writing the paper for the goal of creating a standard definition across different countries.[48]

4.1. Zipf's law

We test Zipf's law by estimating the relationship between the log of population of metropolitan areas and the log of population rank.[49] Within each country/year we estimate:

$$\log\left(\text{Rank}_i - \frac{1}{2}\right) = \alpha - \beta \log(\text{Size}_i) + \varepsilon_i \quad (3)$$

where *i* is a city indicator.

As mentioned, Zipf's law is synonymous with $\beta = 1$. A coefficient greater than 1 suggests that in that particular country the population is more evenly distributed across cities than what the rank size rule predicts (Rosen and Resnick, 1980).[50] The results from estimating the coefficient of Eq. 3 for the same countries studied by Chauvin et al. (2017) are reported on Table 4, which shows estimates using the 15 versions of BEAM in 2000 and 2010. For this exercise we restrict BEAM to the metropolitan areas larger than 100,000 people. We do so for a practical and a conceptual reason. The practical reason is comparability with the results of Chauvin et al. (2017). The conceptual reason is that Zipf's law has been shown to provide a better description of the size distribution of cities in the right tail of the distribution (i.e., for bigger cities). This is because Gibrat's law (which, under certain conditions, implies Zipf's law, Gabaix (1999a); Duranton and Puga (2014)) is more likely to hold in the upper tail of the distribution. As summarized by Gabaix (1999b), "The idea is that new cities are born too far from the upper tail to influence its distribution."[51]

The coefficients estimated for both 2000 and 2010 are smaller in absolute value than those estimated by Chauvin et al. (2017) for each country except China. The ranges of the estimated Zip's law coefficients using the 15 versions of BEAM are quite narrow, even if different assumptions regarding pixel stability and city split yield different sample sizes (see Table A.7). This is reassuring regarding the robustness of our proposed methodology. For the U.S. the range of estimated coefficients is 0.71-0.84 (0.81-0.85) in 2000 (2010). For Brazil the range is 0.82-0.97 (0.84-0.92); for India 0.76-0.95 (0.82-0.96) and for China 0.94-1.06 (0.92-1.00).[52] Our coefficients for Brazil for 2000 are very similar to the one estimated by Soo (2014) for the same year (0.94).

According to these results, China is the country with a size distribution of cities more closely characterized by the rank size rule. Instead, in the U.S., India and Brazil the estimated set of coefficients suggest that the largest cities are too large or too many (or both) relative to what the Zipf's law would predict. In the light of what BEAM measures, this is probably not surprising, as Rosen and Resnick (1980) find that the value of the Pareto exponent is lower for urban agglomerations as compared to cities.[53]

Beyond the comparison with the recent results obtained by Chauvin et al. (2017) for the aforementioned countries, we use BEAM to estimate Eq. 3 for every country in the world. We do so for 2010. The results are summarized in Fig. 3.[54] There is important variation across countries within regions. A few regions have all or most their countries with coefficients consistently larger than −1 (or, consistent with the interpretation above, smaller than 1 in absolute value). This is the case,

[45] Although not reported, our results from this subsection are robust to removing outliers (available upon request).

[46] This is equivalent to saying that the size distribution of cities follows a Pareto distribution with exponent equal to 1.

[47] Gabaix (1999b) shows that if Gibrat's law holds then distribution of city sizes, in equilibrium, will be consistent with Zipf's law.

[48] The data sources used by Chauvin et al. (2017) are the following. For the United States, the 2010 *Consolidated Metropolitan Statistical Areas*, defined by the U.S. Census. For Brazil, *microregions*, composed by agglomerations of contiguous and economically integrated municipalities that have similar economic features, and defined by the Brazilian Institute for Geography and Statistics. For China, administrative *cities*, including provincial-level and prefecture-level areas, which typically comprise both urban and rural territories. For India, *districts*, the second-level administrative division of the country after states and union territories. In the four countries, Chauvin et al. (2017) restrict the samples to urban areas with 100,000 urban dwellers or more.

[49] We follow Gabaix and Ibragimov (2011) and subtract 1/2 from the log of population rank. The authors use simulations to show that this is a better estimate of the coefficient of the power law distribution of city size relative to the traditional approach.

[50] Following this logic, some authors take $\hat{\beta}$ as an indirect indicator of city primacy, as a coefficient smaller than one suggests that one or a few large cities dominate the urban landscape (Soo, 2005). For instance, Rosen and Resnick (1980) show that the correlation between the rank-size rule coefficient and a standard measure of urban primacy (namely the ratio between the size of the largest city to the size of the first *k* largest cities, called '*k*-primacy') is indeed negative. In Table A.6 we compute that correlation using the 15 versions of BEAM and the 5-primacy measure. Reassuringly, and consistent with Rosen and Resnick (1980), we find that both in 2000 and 2010 the large majority of the correlation coefficients are negative.

[51] We also report in Table A.7 the number of metropolitan areas larger than 100,000 people included in each estimation of Eq. 3.

[52] The coefficients estimated by Chauvin et al. (2017) (all for 2010) are 0.91 for the U.S., 1.18 for Brazil, 1.03 for India and 0.91 for China.

[53] Our estimates of Eq. (3) are quite robust to using population raster files other than Landscan to merge with BEAM. Table A.8 of the Appendix compares, for the same four countries, for 2000 and 2010 and using the 15 versions of BEAM, the results using Landscan (column 3), GHS (column 4) and World Pop (column 5). The differences when using Landscan relative than the other two population inputs is in all cases very small and seldom statistically significant (columns 6 and 7). Estimated coefficients of the rank size rule for the entire world, for 2000 and 2010, using the different population sources are available from the authors upon request.

[54] We plot, over the same support, the estimated slope coefficients of the rank-size relationship of each country, together with the 95% confidence intervals. We drop countries with less than 10 cities larger than 100,000 to avoid estimating relationships that are too imprecise. However, below each sub-figure we report the mean and standard deviation of the within region estimated slope coefficient, both for the plotted subsample as well as for the entire sample of countries.

Table 3
Urbanization rate and GDP per capita, all BEAM database.

	Dep var: urbanization rate														
	2000	2010	Panel	2000	2010	Panel	2000	2010	Panel	2000	2010	Panel	2000	2010	Panel
Pixel stability:	0%														
Split threshold:	0 to 1km			0 to 2km			0 to 3 km			0 to 4 km			0 to 5 km		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
log(GDP per capita)	0.1401***	0.1536***	-0.0539*	0.1326***	0.1438***	-0.0487	0.1381***	0.1519***	-0.0489*	0.1394***	0.1528***	-0.0437*	0.1441***	0.1506***	-0.0398
	(0.0257)	(0.0309)	(0.0293)	(0.0185)	(0.0208)	(0.0303)	(0.0213)	(0.0288)	(0.0276)	(0.0213)	(0.0290)	(0.0258)	(0.0202)	(0.0282)	(0.0253)
Observations	157	157	314	155	155	310	154	154	308	153	153	306	151	151	302
R-squared	0.666	0.700	0.984	0.661	0.677	0.981	0.676	0.696	0.982	0.665	0.688	0.982	0.650	0.673	0.982
Dep. variable mean	0.433	0.433	0.433	0.422	0.417	0.419	0.410	0.401	0.406	0.396	0.385	0.390	0.377	0.365	0.371
Dep. variable SD	0.270	0.259	0.264	0.263	0.253	0.257	0.262	0.254	0.257	0.260	0.253	0.256	0.255	0.246	0.250
Indep. variable mean	−4.835	−4.547	−4.691	−4.842	−4.549	−4.696	−4.853	−4.564	−4.708	−4.854	−4.566	−4.710	−4.852	−4.564	−4.708
Indep. variable SD	1.273	1.237	1.262	1.278	1.244	1.268	1.276	1.233	1.261	1.280	1.237	1.265	1.262	1.230	1.252
Pixel stability:	20%														
Split threshold:	0 to 1km			0 to 2km			0 to 3 km			0 to 4 km			0 to 5 km		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
log(GDP per capita)	0.1403***	0.1651***	-0.0542	0.1296***	0.1451***	-0.0357	0.1387***	0.1585***	-0.0287	0.1401***	0.1546***	-0.0085	0.1452***	0.1540***	-0.0243
	(0.0260)	(0.0308)	(0.0346)	(0.0194)	(0.0236)	(0.0419)	(0.0216)	(0.0296)	(0.0349)	(0.0214)	(0.0288)	(0.0357)	(0.0203)	(0.0293)	(0.0379)
Observations	155	155	310	151	151	302	152	152	304	151	151	302	149	149	298
R-squared	0.656	0.689	0.968	0.656	0.652	0.943	0.668	0.689	0.966	0.657	0.680	0.967	0.641	0.674	0.963
Dep. variable mean	0.438	0.445	0.441	0.426	0.424	0.425	0.413	0.420	0.416	0.399	0.408	0.403	0.380	0.395	0.387
Dep. variable SD	0.267	0.270	0.268	0.260	0.266	0.262	0.259	0.262	0.260	0.258	0.260	0.259	0.254	0.256	0.255
Indep. variable mean	−4.806	−4.512	−4.659	−4.801	−4.504	−4.652	−4.823	−4.530	−4.676	−4.824	−4.531	−4.678	−4.822	−4.528	−4.675
Indep. variable SD	1.255	1.208	1.238	1.264	1.218	1.248	1.257	1.204	1.237	1.261	1.208	1.241	1.242	1.200	1.228
Pixel stability:	50%														
Split threshold:	0 to 1km			0 to 2km			0 to 3 km			0 to 4 km			0 to 5 km		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
log(GDP per capita)	0.1391***	0.1577***	-0.0358	0.1334***	0.1474***	-0.0219	0.1371***	0.1534***	-0.0167	0.1377***	0.1529***	-0.0135	0.1421***	0.1566***	-0.0315
	(0.0236)	(0.0282)	(0.0314)	(0.0174)	(0.0204)	(0.0306)	(0.0207)	(0.0288)	(0.0334)	(0.0209)	(0.0286)	(0.0330)	(0.0207)	(0.0285)	(0.0371)
Observations	155	155	310	155	155	310	153	153	306	152	152	304	150	150	300
R-squared	0.669	0.718	0.971	0.659	0.692	0.969	0.667	0.708	0.968	0.656	0.700	0.969	0.639	0.669	0.955
Dep. variable mean	0.428	0.460	0.444	0.418	0.447	0.432	0.403	0.429	0.416	0.389	0.416	0.402	0.376	0.399	0.388
Dep. variable SD	0.267	0.265	0.266	0.261	0.261	0.261	0.259	0.258	0.258	0.257	0.258	0.258	0.254	0.255	0.254
Indep. variable mean	−4.842	−4.549	−4.696	−4.842	−4.549	−4.696	−4.866	−4.574	−4.720	−4.868	−4.576	−4.722	−4.838	−4.553	−4.695
Indep. variable SD	1.278	1.244	1.268	1.278	1.244	1.268	1.268	1.231	1.256	1.272	1.235	1.260	1.253	1.228	1.247

Notes. Robust standard errors clustered at the region-level for cross-sections and at the country-level for panel specifications in parentheses; significance-levels *** 1%, ** 5%, and * 10% refer to two-sided t-tests. Region fixed effects used in cross-section estimates. Country fixed effects used in panel estimates.

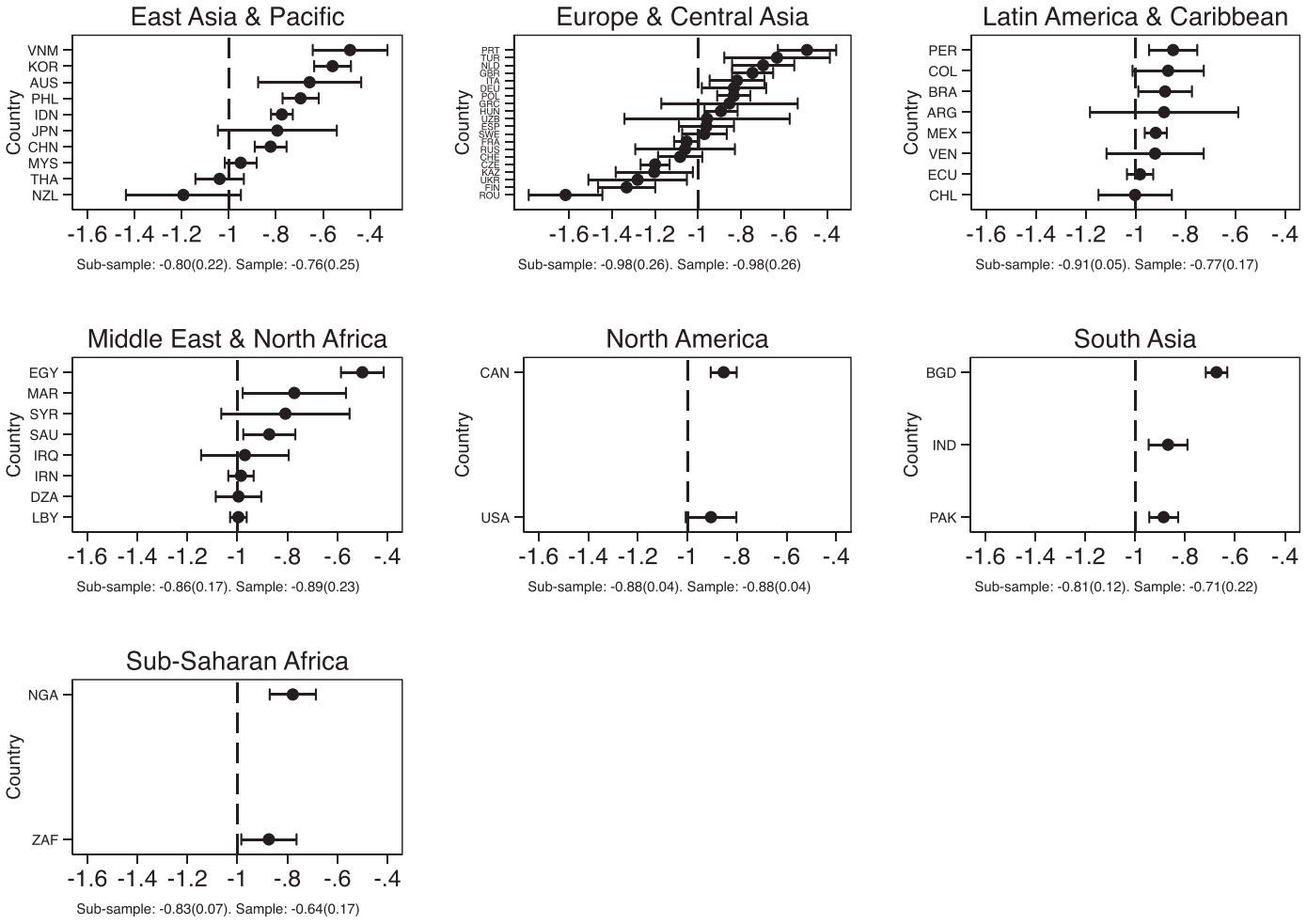

Fig. 3. Zipf's law β estimator by world region in 2010. **Notes:** The graph reports regression point estimates and standard errors based on Gabaix and Ibragimov (2011) following a shifted rank, i.e. rank - 1/2 for each country. The dots show Zipf's coefficient while the horizontal lines represent 95% confidence intervals. We focus on countries with at least 10 cities with more than 100,000 inhabitants in 2010, leaving us with 55 countries. The mean and standard deviation (in parenthesis) are included for both the subsample of 55 countries and the full sample. Countries classified following World Bank's regional classification. Country population is estimated using *BEAM* and Landscan georeferenced population dataset, with a pixel stability threshold of 0% and metropolitan area split threshold of less than 5 km.

for example, of North America, South Asia and Sub-Saharan Africa. The size distribution of cities in several Middle East countries actually meet Zipf's law. This is also the case of a few countries in Latin America and Europe.

Europe is a very interesting region. Our estimates reveal a large variation across the region's countries in terms of the estimated coefficient of the rank size rule test. Many countries have coefficients larger than 1 in absolute value, suggesting that the population in these countries is more evenly distributed than suggest by Zipf's law. Many other have coefficients smaller than 1, consistent with a few large cities concentrating a larger fraction of the population relative to what the law predicts.

Our findings are in line with the results of Soo (2005), who estimates Pareto exponents consistently less than 1. This is also the main finding of Black and Henderson (2003) for the case of the U.S.. This may be explained by the increasing sub-urbanization experienced by countries in the past few decades. Our estimates of Gibrat's law are also consistent with these findings.

4.2. Gibrat's law

Gibrat's law states that the growth of urban areas is independent from its size. If bigger cities grew faster then they would tend to concentrate a larger share of population and economic activity over time, becoming a sort of urban black holes. Gibrat's law can be tested by estimating, at the city level, the relationship between the change in population over a specific period and the initial population level. If the estimated correlation coefficient is not statistically distinguishable from zero this would be consistent with Gibrat's law.[55] Specifically, for each country we estimate:

$$\log(\text{Pop. Growth})_i^{2000-2010} = \alpha + \beta \log(\text{Population})_i^{2000} + \varepsilon_i \quad (4)$$

where i is the city indicator, with $\log(\text{PopGrowth})_i^{2000-2010}$ the change in the logarithm of city population from 2000 to 2010.

Using BEAM, Table 5 reports the estimated slope coefficients, focusing on the countries analyzed by Chauvin et al. (2017) and using the 15 versions of our data.[56] The estimated coefficient is negative in most cases and significant for the U.S. and India in all the specifications, and for China in a third of the specifications. For

[55] One caveat is that, in the presence of measurement error, city population growth may reverse to the mean Gabaix and Ioannides (2004).

[56] To estimate Eq. 4, we keep a balanced sample of cities, dropping from the estimation cities that were smaller than 100,000 in 2000 (and thus were not included in the dataset) but larger than the threshold in 2010 (thus making it into the dataset).

Table 4
Zipf's coefficient, by country.

2000					
Splitting threshold	0 to 1km	0 to 2km	0 to 3km	0 to 4km	0 to 5km
Pixel stability	(1)	(2)	(3)	(4)	(5)
	Brazil				
0%	−0.86***	−0.86***	−0.94***	−0.96***	−0.97***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
20%	−0.89***	−0.89***	−0.82***	−0.91***	−0.92***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
50%	−0.85***	−0.86***	−0.86***	−0.94***	−0.94***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
	China				
0%	−1.06***	−1.06***	−1.05***	−1.04***	−1.02***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
20%	−0.97***	−0.96***	−0.95***	−0.95***	−0.94***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
50%	−1.00***	−1.01***	−1.00***	−0.98***	−0.96***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
	India				
0%	−0.88***	−0.92***	−0.95***	−0.93***	−0.92***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
20%	−0.76***	−0.81***	−0.83***	−0.85***	−0.84***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
50%	−0.80***	−0.85***	−0.88***	−0.90***	−0.89***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
	United States				
0%	−0.74***	−0.79***	−0.81***	−0.83***	−0.84***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
20%	−0.74***	−0.77***	−0.75***	−0.77***	−0.83***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
50%	−0.71***	−0.77***	−0.77***	−0.80***	−0.80***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
2010					
Splitting threshold	0 to 1km	0 to 2km	0 to 3km	0 to 4km	0 to 5km
Pixel stability	(1)	(2)	(3)	(4)	(5)
	Brazil				
0%	−0.89***	−0.90***	−0.92***	−0.92***	−0.92***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
20%	−0.84***	−0.84***	−0.86***	−0.86***	−0.87***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
50%	−0.89***	−0.90***	−0.91***	−0.91***	−0.91***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
	China				
0%	−1.00***	−0.98***	−0.98***	−0.97***	−0.95***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
20%	−0.93***	−0.92***	−0.93***	−0.94***	−0.93***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
50%	−0.98***	−0.98***	−0.94***	−0.94***	−0.94***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
	India				
0%	−0.96***	−0.95***	−0.93***	−0.91***	−0.89***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
20%	−0.82***	−0.83***	−0.85***	−0.83***	−0.83***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
50%	−0.90***	−0.90***	−0.90***	−0.89***	−0.88***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
	United States				
0%	−0.84***	−0.84***	−0.84***	−0.84***	−0.85***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
20%	−0.81***	−0.82***	−0.83***	−0.82***	−0.83***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
50%	−0.81***	−0.82***	−0.83***	−0.82***	−0.83***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)

Notes. Columns report the point estimate of the regression of (log) population on the (log) rank-1/2 -following Gabaix and Ibragimov (2011)- of cities with at least 100,000 inhabitants. The population is estimated using *BEAM* and Landscan georeferenced population dataset with a pixel stability threshold of 0%, 20% or 50% and city split threshold of less than 1 km, 2 km, 3 km, 4 km or 5 km. Robust standard errors are shown in parentheses: *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level.

Brazil is only significant in two out of the 15 specifications, and given the amount of large cities in the country the lack of statistical significance is probably not associated with lack of power. This implies that, if anything, Gibrat's law applies only for the case of Brazil.

The results for the U.S. and India suggest that, over the decade-long period, initially smaller cities grew faster in population relative to initially larger cities. This is consistent with our test of Zipf's law for these two countries, as we found that the size rank rule coefficient is unequivocally bigger in 2010 than in 2000.

Table 5
Gibrat's law by country.

Splitting threshold	Brazil				
Pixel stability	0 to 1km (1)	0 to 2km (2)	0 to 3km (3)	0 to 4km (4)	0 to 5km (5)
0%	−0.32**	−0.33**	−0.16	0.03	0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
20%	0.01	0.00	−0.16	0.01	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
50%	−0.22	−0.22	−0.24*	−0.14	−0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	China				
0%	−0.19***	−0.02	−0.06	−0.10	−0.08
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
20%	−0.02	−0.01	−0.01	−0.04**	−0.06***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
50%	−0.00	0.01	0.01	−0.02	−0.09***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	India				
0%	−0.46***	−0.23***	−0.18***	−0.06	−0.10
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
20%	−0.24***	−0.16**	−0.17**	−0.10***	−0.09***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
50%	−0.29***	−0.19***	−0.17**	−0.12***	−0.08***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	United States				
0%	−0.31***	−0.37***	−0.26***	−0.17	−0.25***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
20%	−0.23***	−0.26***	−0.27***	−0.18***	−0.15**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
50%	−0.27***	−0.27***	−0.32***	−0.14**	−0.14**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes. Columns report the point estimates on the regression of the change in the logarithm of city population from 2000 to 2010 of cities with at least 100,000 inhabitants on the logarithm of their population in 2000. Population is estimated using *BEAM* and Landscan georeferenced population dataset with a pixel stability threshold of 0%, 20% or 50% and a city split threshold of less than 1 km, 2 km, 3 km, 4 km or 5 km. Robust standard errors are shown in parentheses: *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level.

A potential concern is that the different degree to which the presence of clouds hides the extent of metropolitan areas (by covering night lights that can be detected by satellite images), may bias these results. Thus, for robustness, we estimate Eq. (4) controlling for geo-localized cloud coverage. Table A.10 of the Appendix presents point estimates using the two specifications (with and without this control) as well as *t*-test of the difference between the two. By and large, there are no significant differences when the presence of clouds is taken into account. We carry out the same robustness test when estimating Eq. (3) on Zipf's coefficient in Table A.9 in the Appendix. As with Gibrat's coefficient, we do not find a difference in Zipf's coefficient with and without controlling for cloud cover.

5. Discussion: Potential disadvantages and further robustness

It is important to acknowledge some disadvantages of nighttime light data that might affect the estimation of the size of metropolitan areas. In particular, in this section we consider three important caveats: the potential bias coming from the income gradient in energy consumption; the potential bias coming from different parts of the world being exposed to more clouds (which may affect the capacity of satellite pictures to identify ground activity); and the potential overestimation coming from the imagery resolution.

5.1. Energy consumption and nighttime lights

We have argued that the procedure described in Section 2.3.2 deals with potential measurement error derived from using traditional definitions of the extent of a city. There are, however, other potential sources of measurement error that are specific to our nighttime light input. Specifically, the income gradient in energy consumption may affect our capacity to identify cities in relatively poor areas of the world.

Our response to this potential criticism is twofold. First, recall that we do not impose a DN threshold in our methodology but rather use nighttime luminosity in the *extensive margin* to identify cities. Second, we explore the extent to which country level variation in per capita energy consumption is correlated with the size of metropolitan areas of at least 50,000 people as measured by BEAM. To that end, we estimate the correlation between energy consumption in (log) kWh per capita and the (log) size of metropolitan areas aggregated at the country level using the 15 versions of BEAM. We also control for country and time fixed effects, as well as for GDP per capita and the urbanization rate.

The results are reported in Table A.2 and suggest that the correlation between these two variables is not significantly different from zero. We reckon that this provides suggestive (albeit indirect) evidence against the possibility that differential levels of energy consumption across different levels of development substantially bias our measure of city size.

5.2. Clouds and nighttime light

Atmospheric conditions may introduce spatial bias, especially for optical raster images such as nighttime light data. For instance, extreme cloud cover –as well as dust and haze– may introduce blurring (as clouds scatter light across space when light is too bright) but also may block light from satellite lenses. According to NOAA, the former is addressed in DMSP-OLS Nighttime Lights by generating cloud-free "stable lights". However, clouds (or other atmospheric phenomena) might still interfere with satellites optical sensors.[57]

In addition to addressing any potential blurring following Abrahams et al. (2018) and using for robustness a range of different pixel stability cut-offs, in this section we provide additional evidence that the presence of clouds is unlikely to significantly affect our estimates of the size distribution of cities. In particular, Tables A.9 and A.10 of the Appendix compare the coefficients obtained from estimating Eqs. (3) and (4) with and without controlling for cloud coverage. We do so for the four countries studied by Chauvin et al. (2017), for both 2000 and 2010 and for the 15 versions of BEAM computed for this paper. With the exception of Zipf's coefficient for China in 2000 (and when the stability threshold is either 20% or 50%), the estimated slope coefficients are statistically indistinguishable from one another with and without controlling for average cloud coverage. This suggests that atmospheric conditions are unlikely to introduce an important bias in using BEAM to study the size distribution of cities, as well as the relationship between city size and growth.[58]

5.3. Pixel resolution and potential overestimation

Even after applying all the correction procedures discussed in Section 2.3.2, BEAM likely still overestimates the size of cities due to the low spatial resolution of the input satellite pictures. Indeed, due to limitations of OLS sensors, nighttime light composites have a resolution of 1 km × 1 km per cell, which may decrease our capacity to accurately identify the peripheral limits of metropolitan areas, especially compared to methodologies that use higher resolution images of

[57] For instance, Abrahams et al. (2018) apply NOAA's cloud-free stable lights to study the size of 15 cities and find an overestimation of 49%. This motivates the authors' methodology for introducing the pixel-stability threshold that we use in this paper (in fact they call it "cloud free" night threshold). However, the authors note that this procedure reduces the overestimation to 9%, which is a substantial but only partial reduction.

[58] Moreover, recall from the comparison of Appendix Figs. A.1 and A.2 with Fig. A.3 that world areas with more cloud coverage are not the same in which BEAM fails to identify cities listed by AUE or GHS.

urban built-up.[59] To be sure, more recent nightlight files (produced by *Suomi National Polar-orbiting Partnership satellite's Visible Infrared Imaging Radiometer Suite*, Suomi NPP/VIIRS) do have higher spatial resolution. However, to the best of our knowledge most of the procedures available to clean DMSP-OLS files (such as bright-light saturation or in-flight calibration) have not yet been developed (Baugh et al., 2013). Moreover, VIIRS coverage starts only in 2011. As mentioned, our results are largely robust to applying a pixel stability threshold of 50%, which likely reduces the scope for overestimation.[60]

A second major disadvantage of DMPS OLS is its unavailability after 2013. However, recent studies show that it is possible to simulate DMSP-OLS composites using VIIRS. For instance, Li et al. (2017) propose an inter calibration model to carry such simulation for Syria, thus proving an interesting avenue of future researchers interested in studying long-term changes of urban footprint using high-resolution clean data.

6. Conclusion

Data-driven spatial definitions of the size and growth of urban areas face a trade-off between the precision of estimates and the empirical practicality of the estimation process, mainly given by the availability of the required inputs and by computational demand. In general, highly precise measures imply low empirical practicality. The measure and dataset introduced in this paper offer a highly precise estimate (at least when compared to other databases) that is also computationally tractable, as well as transparent and replicable. Our replication data is available to any researcher.

Particularly, we use nighttime light satellite data, that we pre-process using various existing correction procedures, to estimate the size of urban areas in the entire world. The result is a comprehensive dataset of global metropolitan areas greater than 50,000 people in two time periods, 2000 and 2010. We also merge our data with existing georeferenced population datasets to compute various comparable city-level measures, including the size of cities, urban population and population density. This also allows us to construct country-level urbanization rates, a variable that is hotly debated in the urban economics literature precisely because it is generally computed using country-specific administrative definitions of urban boundaries (see Roberts et al., 2016).

Because our measures are comparable for cities of all the world's regions, and are thus independent of the level of development and of administrative definitions, the dataset allows us to revisit some of the basic stylized facts about urbanization in the world. For instance, we compute statistics that describe the size distribution of cities with countries as well as the relationship between city size and growth. Particularly, we test Zipf's law and Gibrat's law and generally find no support for either.

Our results do not necessarily support previous findings regarding global stylized facts, but at the same time they are quite robust to changes in the arbitrary parameters that are necessary to estimate our dataset, in particular the stability of lit pixels and the size of urbanized corridors that connect large metropolitan areas. Nevertheless, we discuss a few potential limitations of our measure and point to possible corrections that are however more computationally intensive and require ancillary data. The future of remote sensing tools and nighttime light appears to be promising as an increasing number of satellites become available to follow human migration and urban built-up patterns. DMSP nighttime light data (our main data source) comes with several measurement pitfalls, but recently launched satellites with new sensors are incorporating new technology, including multi-spectral bands to capture land use, radiometrically calibrated images to solve light saturation problems to analyze highly luminous areas in a reliable way, and higher levels of spatial resolution to increase the accuracy of delineating urban expansion, size and shape. The urban economics literature is already moving towards exploiting these type of datasets, that permit better cross- and within- country comparable analyses. BEAM is an example of such datasets.

[59] For instance, MODIS, Landsat and TerraSAR/TanSAR provide higher resolution images of respectively 500 m^2, 38 m^2, and 12 m^2 per cell.

[60] As mentioned on footnote [57], applying a 20% threshold Abrahams et al. (2018) reduce the overestimation of the uncorrected, blurred DMSP-OLS images from 49% to 9%. Overestimation can further be reduced by correcting BEAM estimates with additional information, for instance GHS measures of urban footprint. Our results are robust to this approach, and the results are available from the authors upon request.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jue.2020.103254.

CRediT authorship contribution statement

Rafael Ch: Conceptualization, Methodology, Software, Data curation, Writing - original draft. **Diego A. Martin:** Software, Data curation. **Juan F. Vargas:** Supervision, Conceptualization, Methodology, Writing - original draft, Writing - review & editing.

References

Abrahams, A., Lozano-Gracia, N., Oram, C., 2018. Deblurring DMSP nighttime lights: a new method using Gaussian filters and frequencies of illumination. Remote Sens. Environ. 210 (1), 242–258.
Acemoglu, D., Johnson, S., Robinson, J.A., 2002. Reversal of fortune: geography and institutions in the making of the modern world income distribution. Q. J. Econ. 117, 1231–1294.
Angel, S., Blei, A.M., Civco, D.L., Parent, J., 2012. Atlas of urban expansion. Lincoln Institute of Land Policy Cambridge, MA.
Angel, S., Sheppard, S.C., Civco, D.L., 2005. The dynamics of global urban expansion. Transp. Urban Dev. Dep. World Bank.
Bairoch, P., 1991. Cities and Economic Development: From the Dawn of History to the Present. University of Chicago Press.
Baugh, K., Elvidge, C.D., Hsu, F.-C., Zhizhin, M., 2013. Why VIIRS data are superior to DMSP for mappping nighttime lights. Proc. Asia-Pacif. Adv. Netw. 35, 62–69.
Black, D., Henderson, V., 2003. Urban evolution in the USA. J. Econ. Geograph. 3 (4), 343–372.
Bleakley, H., Lin, J., 2012. Portage and path dependence. Q. J. Econ. 127, 587–644.
Burchfield, M., Overman, H.G., Puga, D., Turner, M.A., 2006. Causes of sprawl: a portrait from space. Q. J. Econ. 121 (2), 587–633.
Chauvin, J.P., Glaeser, E., Ma, Y., Tobio, K., 2017. What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. J. Urban Econ. 98, 17–49.
Combes, P.-P., Duranton, G.D., Gobillon, L., 2019. The costs of agglomeration: house and land prices in French cities. Rev. Econ. Stud. 86 (4), 1556–1589.
de Sherbinin, A., Adamo, S., 2015. Ciesin'S experience in mapping population and poverty. United Nations Expert Group Meeting on Strengthening the Demographic Evidence Base for the Post-2015 Development Agenda, Population Division, Department of Economic and Social Affairs.
Dijkstra, L., Poelman, H., 2014. A harmonised definition of cities and rural areas: the new degree of urbanization. Regional Working Paper, European Commission, Directorate–General for Regional and Urban Policy.
Dittmar, J.E., 2011. Information technology and economic change: the impact of the printing press. Q. J. Econ. 126 (3), 1133–1172.
Donaldson, D., Storeygard, A., 2016. The view from above: applications of satellite data in economics. J. Econ. Perspect. 30 (4), 171–198.
Doxsey-Whitfield, E., MacManus, K., Adamo, S., Pistolesi, L., Squires, J., Borkovska, O., Baptista, S., 2015. Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4 (GPWv4). Paper. Appl. Geogr. 1 (3), 226–234.
Duranton, G., 2015. A proposal to delineate metropolitan areas in Colombia. Revista Desarrollo y Sociedad (75) 223–264.
Duranton, G., Puga, D., 2014. The growth of cities. Handbook of Economic Growth, 2. Elsevier, pp. 781–853.
Elvidge, C., Baugh, K., Kihn, E., Kroehl, H., Davis, E., 1997. Mapping city lights with night-time data from the DMSP operational linescan system. Photogramm. Eng. Remote Sens. 63, 727–734.
Elvidge, C., Baugh, K., Tilottama, G., Zhizhin, M., 2009. A fifteen year record of global natural gas flaring derived from satellite data. Energies 2, 595–622.
Elvidge, C.D., 2000. Radiance calibration of DMSP-OLS low-light imaging data of human settlements (cd-rom). US Department of Commerce, National Oceanographic and Atmospheric Administration.
Feenstra, R.C., Inklaar, R., Timmer, M.P., 2015. The next generation of the Penn World Table. Am. Econ. Rev, 105 (10), 3150–3182.
Gabaix, X., 1999. Zipf'S law and the growth of cities. Am. Econ. Rev. 89 (2), 129–132.
Gabaix, X., 1999. Zipf'S law for cities: an explanation. Q. J. Econ. 114 (3), 739–767.

Gabaix, X., Ibragimov, R., 2011. Rank- 1/2: a simple way to improve the ols estimation of tail exponents. J. Bus. Econ. Stat. 29 (1), 24–39.
Gabaix, X., Ioannides, Y.M., 2004. The evolution of city size distributions. In: Handbook of Regional and Urban Economics, 4. Elsevier, pp. 2341–2378.
Goldblatt Ran, S.M.F., Tellman, B., Clinton, N., Hanson, G., Georgescu, M., Wang, C., Serrano-Candela, F., Khandelwal, A., Cheng, W.-H., Balling Jr, R.C., 2018. Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. Remote Sens. Environ. 205 (253–275).
Harari, M., 2017. Cities in bad shape: urban geometry in India. Working paper, The Wharton School.
Henderson, M., Yeh, E., Gong, P., Elvidge, C.D., Baugh, K., 2003. Validation of urban boundaries derived from global night-time satellite imagery. Int. J. Remote Sens. 24 (3), 595–609.
He, C., Shi, P., Li, J., Chen, J., Pan, Y., Li, J., Zhuo, L., Ichinose, T., 2006. Restoring urbanization process in china in the 1990s by using non-radiance-calibrated DMSP/OLS nighttime light imagery and statistical data. Chin. Sci. Bull. 51 (13), 1614–1620.
Henderson, V., Storeygard, A., Weil, D.N., 2011. A bright idea for measuring economic growth. Am. Econ. Rev. 101 (3), 194–199.
Hsu, F.-C., Baugh, K., Ghosh, T., Zhizhin, M., 2015. DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. Remote Sens. Environ. 7, 1855–1876.
Imhoff, M., Lawrence, W., Stutzer, D., Elvidge, C., 1997. A technique for using composite DMSP/OLS "city lights" satellite data to map urban area. Remote Ses. Environ. 61, 361–370.
Ioannides, Y., Skouras, S., 2013. US city size distribution: robustly pareto, but only in the tail. J. Urban Econ. 73 (1), 18–29.
Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., Ermon, S., 2016. Combining satellite imagery and machine learning to predict poverty. Science 353 (6301), 790–794.
Jedwab, R., Vollrath, D., 2019. The urban mortality transition and poor-country urbanization. Am. Econ. J. 11 (1), 223–275.
Kuznets, S., 1968. Toward a Theory of Economic Growth. W.W. Norton & Company.
Li, X., Xu, H., Wu, C., 2017. Intercalibration between DMSP/OLS and VIIRS night-time light images to evaluate city light dynamics of Syria's major human settlement during Syrian civil war. Int. J. Remote Sens. 38 (21), 5934–5951.
Liu, Z., He, C., Zhang, Q., Yang, Y., 2012. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. Landsc. Urban Plan. 106, 62–72.
Lloyd, C., Sorichetta, A., Tatem, A., 2017. High resolution global gridded data for use in population studies. Sci. Data 4 (170001).
Lowe, M., 2014. The privatization of African rail. Working paper.
Martinez, L.R., 2018. How much should we trust the dictator's GDP estimates? Working paper, Harris School of Public Policy.
Michalopoulos, S., Papaioannou, E., 2013. Pre-colonial ethnic institutions and contemporary African development Econometrica 81, 113–152.
OECD, 2012. Redefining "Urban": A New Way to Measure Metropolitan Areas. Paris: OECD.
OECD, 2013. Definition of functional urban areas (fua) for the OECD metropolitan database. OECD Working paper.
Roberts, M., Blankespoor, B., Deuskar, C., Stewart, B., 2016. Urbanization and development: is Latin America and the Caribbean different from the rest of the world? Working paper, The World Bank.
Rosen, K.T., Resnick, M., 1980. The size distribution of cities: an examination of the Pareto law and primacy. J. Urban. Econ. 8 (2), 165–186.
Rosenthal, S.S., Strange, W.C., 2004. Evidence on the nature and sources of agglomeration economies. Handbook Reg. Urban Econ. 4, 2119–2171.
Saiz, A., 2010. The geographic determinants of housing supply. Q. J. Econ. 125 (3), 1253–1296.
Soo, K.T., 2005. Zipf's law for cities: a cross-country investigation. Reg. Sci. Urban Econ. 35 (3), 239–263.
Soo, K.T., 2014. Zipf, Gibrat and geography: evidence from China, India and Brazil. Pap. Region. Sci. 91 (1), 159–182.
Storeygard, A., 2016. Farther on down the road: transport costs, trade and urban growth in Sub-Saharan Africa. Rev. Econ. Stud. 83 (3), 1263–1295.
Sutton, P., Roberts, D., Elvidge, C.D., 1997. A comparison of night- time satellite imagery and population density for the continental United States. Photogramm Eng Remote Sens. 63, 1303–1313.
Sutton, P., Roberts, D., Elvidge, C., Miej, H., 1999. Census from heaven: an estimate of the global human population using night-time satellite imagery.. Paper presented at the Western Regional Science Association annual meeting, Ojai, California, USA,.
Uchida, H., Nelson, A., 2008. Agglomeration index: towards a new measure of urban concentration. United Nations, Working Paper.
Weidmann, N.B., Schutte, S., 2016. Using night light emissions for the prediction of local wealth. J. Peace Res. 54 (2), 1–16.
Wu, J., He, S., Peng, J., Li, W., Zhong, X., 2013. Intercalibration of DMSP-OLS night-time light data by the invariant region method. Int. J. Remote Sens. 34 (20), 7356–7368.
Zhang, Q., Seto, K., 2011. Mapping urbanization dynamics at regional and global scales using multip-temporal DMSP/OLS nighttime light data. Remote Ses. Environ. 115 (9), 2320–2329.
Zhao, N., Zhou, Y., Samson, E.L., 2015. Correcting incompatible dn values and geometric errors in nighttime lights time-series images. IEEE Trans. Geosci. Remote Sens. 53 (4), 2039–2049.