

UNIVERSITY OF WAIKATO

**Hamilton
New Zealand**

**Remotely (and wrongly) too equal:
Popular night-time lights data understate spatial inequality**

Xiaoxuan Zhang, John Gibson and Xiangzheng Deng

Working Paper in Economics 13/22

November 2022

Xiaoxuan Zhang
University of Waikato

Corresponding Author

John Gibson
School of Accounting, Finance
and Economics
University of Waikato
Private Bag 3105
Hamilton, 3240
New Zealand

Tel: +64 (7) 838 4289
Email: jkgibson@waikato.ac.nz

Xiangzheng Deng
Institute of Geographic Science and Natural Resources Research
Chinese Academy of Sciences

Abstract

Several studies in economics and regional science use Defense Meteorological Satellite Program (DMSP) night-time lights data to measure spatial inequality. These DMSP data are a poor proxy in this context because they have spatially mean-reverting errors, yielding significantly lower inequality estimates than what sub-national GDP data show. Inequality estimates from DMSP are also lower than what newer, research-focused and more accurate, satellites show from their observations of the earth at night. In this paper, county-level data from the United States and China are used to demonstrate the understatement of spatial inequality when DMSP data are used. In both settings, benchmark data on sub-national GDP are available for establishing the level and trend in spatial inequality, which is then used to assess the accuracy of the estimates coming from remote sensing sources. In the rush to use big data it is important to not lose sight of basic measurement error features of some of these data sources.

Keywords

DMSP

mean-reverting error

night lights

spatial inequality

VIIRS

JEL Classification

E01

R12

Acknowledgements

The financial support from Marsden Fund project UOW1901 and the assistance from Chao Li and Geua Boe-Gibson is gratefully acknowledged.

I. Introduction

A growing literature uses data on satellite-detected night-time lights to study spatial inequality. Our review found over 25 papers in this literature, a majority of which have published in the last four years. Potential advantages of this remote sensing approach include greater comparability over space irrespective of countries' varying levels of statistical capacity, cheaper and more timely data, and the opportunity to extend inequality estimates to spatial units below the level at which GDP data or survey data are reported.¹ In contrast, the existing situation with traditional approaches is aptly described by Mirza et al (2021:1) as one where “our knowledge about inequality in the developing world is limited by paucity, poor quality, uncertainty, and incomparability of data.”

There is a potential flaw in this new research on inequality, because most studies rely on using Defense Meteorological Satellite Program (DMSP) data. This involves repurposing images originally gathered to detect clouds for short-term Air Force weather forecasts. While a seminal study in economics finds that these DMSP data can also be used for long-run observations of economic activity on earth (Henderson et al, 2012), that study only required that the measurement errors in DMSP data are independent of the errors in reported GDP (because an optimal mix of the two types of noisy data was used to better proxy for the true but unknown economic activity). In contrast, the recent inequality studies use the DMSP data directly and because these data have spatially mean-reverting errors (Gibson, 2021) they are singularly ill-suited for measuring spatial inequality, even if they are a reasonable proxy in the context where Henderson et al (2012) used them (for studying growth in national GDP).

A consequence of these spatially mean-reverting errors is that DMSP data yield far lower inequality estimates than what sub-national GDP data show. The inequality estimates from DMSP are also lower than what newer, research-focused and more accurate, satellites show from their observations of the earth at night. Moreover, our results here suggest that not just the apparent level of spatial inequality, but also the temporal trend, may be affected if the DMSP data are used. In other words, even though using night-time lights data has potential to overcome some constraints on our understanding of inequality, the data that have been most commonly used to date could introduce some new misunderstandings.

In the next section we discuss the sources of spatially mean-reverting errors in DMSP data. A visual example is given in Section III, to show how DMSP blurring and top-coding distort apparent patterns of spatial variation in economic activity. Our literature review in Section IV shows that recent inequality studies ignore these spatially mean-reverting errors. Section V has county-level evidence from the United States and China on the understatement of spatial inequality when the DMSP data are used. Section VI has our conclusions.

¹ In comparison, survey-based approaches may be incompatible over time and space, with measured inequality varying with survey welfare indicators (Deininger and Squire, 1996) and reference periods (Gibson et al, 2003). Surveys are also expensive, at up to US\$4000 per household, are fielded infrequently, and can be slow to release data (Sharp et al, 2022). Moreover, samples are often just large enough for results to be representative at the first sub-national level (e.g. provinces); survey-to-census imputation (e.g. Elbers et al, 2003) lets inequality estimates be formed at lower levels (e.g. counties) but limits temporal coverage because of census infrequency. The other traditional source of data for estimating spatial inequality is national and regional accounts data. Wide variation in statistical capacity and data availability impose limits on the usefulness of these data (Henderson et al, 2012).

II. The Spatially Mean-Reverting Errors in DMSP Night Lights Data

The spatially mean-reverting errors in DMSP data have two main sources, and both reflect some of the inherent limitations of DMSP sensors and data management (Abrahams et al, 2018). First, DMSP images are blurred, so some light is attributed to places from where it was not emitted. Earlier remote sensing studies used terms like ‘blooming’ and ‘overflow’ to describe aspects of this blurring (e.g. Small et al, 2005) but that discussion did not pinpoint sources of the problem. Moreover, when economists started to use the night-time lights data they focused on just one aspect of the earlier literature—that environmental factors such as snow and water may cause reflections and create overflow—and inferred that the blurring problem should therefore matter in only a few places (e.g. Michalopoulos and Papaioannou, 2014). In fact, blurring is a universal feature of DMSP data that shows up everywhere.

The blurring occurs for at least three reasons. First, away from the nadir of the 3000 kilometre (km) wide scan the DMSP sensor views the earth at an angle and this causes the Field-of-View (FoV) to expand (by 400% at the scan edge).² All light from the expanded FoV gets (wrongly) attributed to a far smaller pixel in the centre of the FoV. To limit the bias the National Oceanic and Atmospheric Administration (NOAA) only use pixels between the nadir and the half-scan (where the FoV is expanded by just 240%) when processing DMSP images but this does not reverse the blurring. Second, the on-board computers cannot hold all the data, so to conserve memory the pixels are aggregated to 5×5 blocks. Third, random geo-location errors, with a mean of about 3 km, also spread recorded light away from its point of origin (Tuttle et al, 2013). A method to de-blur DMSP images was developed by Abrahams et al (2018) but our review found no inequality studies using this method with DMSP data.

The net effect of the expanded FoV, the pixel aggregation, and the geo-location errors is that the ground footprint for the DMSP sensor is 25 km² at nadir (Elvidge et al, 2013). At the edge of the half-scan it may be up to 60 km² because the FoV expands when viewing the earth at an angle for off-nadir pixels. This spatial resolution is far coarser than the footprint of the sensors used by newer, research-orientated, night lights data sources such as VIIRS Night Lights (VNL) and NASA’s Black Marble (BM).³ A comparison of various night lights data sources in Table 1 shows the sensors used by VNL and BM have constant 0.55 km² footprint, where this invariance is achieved by turning off some of the detector elements away from the nadir to counteract the effect of viewing the earth at an angle. These newer data sources are at least 45-times more spatially precise than are the DMSP images (Elvidge et al, 2013).

The 25 km² (or larger) ground footprint of the DMSP sensor is also far larger than the 30 arc second (roughly 0.9km × 0.9km at the equator or 0.9km × 0.6km at 45° latitude) grid onto which DMSP data are allocated. The inequality literature using DMSP data seems to not recognize the distinction between the sensor resolution and the scale of the output grid. For example, a recent study simply states that DMSP has “spatial resolution of 30 × 30 arc seconds

² To provide intuition for this effect, if a flashlight shines directly down onto the ground it illuminates a circle but if it shines on a spot on the ground somewhere ahead it illuminates an ellipse that is far larger than the circle.

³ Technical details on these two data sources, which both rely on the Suomi/NPP satellite but process the data in different ways, can be found in Elvidge et al (2021) and Román et al (2018).

(approximately 1×1 km)” (Sangkasem & Puttanapong, 2022: 830). In fact, spatial resolution of the sensor is far coarser than the output grid, so differences between each of the 30 arc second DMSP output pixels are blurred. The threat to spatial inequality estimates is easiest to explain with the pixel aggregation to conserve on-board memory. Consider an area with one brightly-lit pixel surrounded by 24 unlit pixels, which would suggest a high degree of local inequality. With pixel aggregation, luminosity attributed to the brightly-lit pixel will be dragged down to the mean for the 25 pixels, while for the 24 unlit pixels their values will be overstated; aggregation is inherently mean-reverting. The impact of this bias will depend on the spatial scale of an analysis; it should matter more for smaller spatial units than for very large ones such as countries or regions at the first sub-national level.

Table 1: Attributes of DMSP and other night-time lights data sources

	DMSP	VNL	BlackMarble
<i>Satellite/Sensor Attributes</i>			
Operator	US DoD		NASA/NOAA
Original purpose	Detect moon-lit clouds for weather forecasts		Earth observation for scientific research
Available years	1992-2021		2012-2021
Ground footprint at nadir	25 km ²		0.55 km ²
Scan width	3000 km (but only the center half is processed)		3000 km
Revisit time	12 hr		12 hr
Pixel saturation	Saturated		Not saturated
On-board calibration	No		Yes
<i>Data Products</i>			
Creator of annual composites	EOG	EOG	NASA
Spatial resolution of processed data	30 arc second grid	15 arc second grid	15 arc second grid
Quantization	6-bit ($n=64$)	14 bit ($n=16,384$)	16 bit ($n=65,536$)
Masking of ephemeral light sources	No	Yes	Yes
Stray-light correction	No	Yes, from 2014	Yes
User control over angle of detection	No	No	Yes

Note: DoD is Department of Defense, EOG is Earth Observation Group

The second source of mean-reverting error in DMSP data is top-coding, due to two factors. First, DMSP was developed to measure clouds rather than to measure lights on earth, and unrecorded sensor amplification occurs during the dark part of the lunar cycle when the cloud-tops are no longer visible in raw moonlight (Hsu et al, 2015). Given that the sensor has only a low dynamic range, when amplification is increased the images for brightly lit parts of the earth, such as central business districts, are saturated with light. Second, pixel aggregation to save memory sees the original 8-bit values divided by four and top-censored at 63, to give the widely used 6-bit DN (digital number). Top-coding shows up in anomalous patterns, such as key infrastructure like Heathrow airport (the most brightly lit feature in England) seeming to be no brighter than the surrounding area (Gibson, 2021) and city centers seeming to be no brighter than lower density suburbs (Bluhm and Krause, 2022).

In contrast to DMSP, nightly images used by the VNL and BM data are not subject to saturation so there is no top-coding problem with these data. The sensors for these newer data sources have far wider dynamic range, covering seven orders of magnitude of radiance (while DMSP only covers two orders of magnitude) so they can simultaneously observe dimly and bright lit areas. Moreover, when the signal is quantized it is with 14-bit resolution (VNL) or 16-bit resolution (BM), providing up to 65,536 different values as opposed to the coarse 6-bit resolution of DMSP data with only 64 possible values (Table 1). The VNL and BM data also have inbuilt masking of ephemeral light sources and correction for stray lights (such as from solar glare in long summer evenings) and users have some control over the angle of detection by restricting attention to either near-nadir or off-nadir pixels (BM only).

The improved quality of VNL and BM data, versus the relatively crude DMSP data, shows up when formal tests for mean-reverting error are conducted. The econometric basis for these tests (Black et al, 2000; Gibson and Kim, 2010) relates observed values of a variable of interest, y^* to the presumed true values, y by:

$$y^* = \theta + \lambda y + v \quad (1)$$

Textbook classical measurement error assumes $\theta = 0, \lambda = 1$ and $E(v) = cov(y, v) = 0$, so that just white noise is added to the true value. In contrast, if the measurement errors are mean-reverting, $0 < \lambda < 1$.

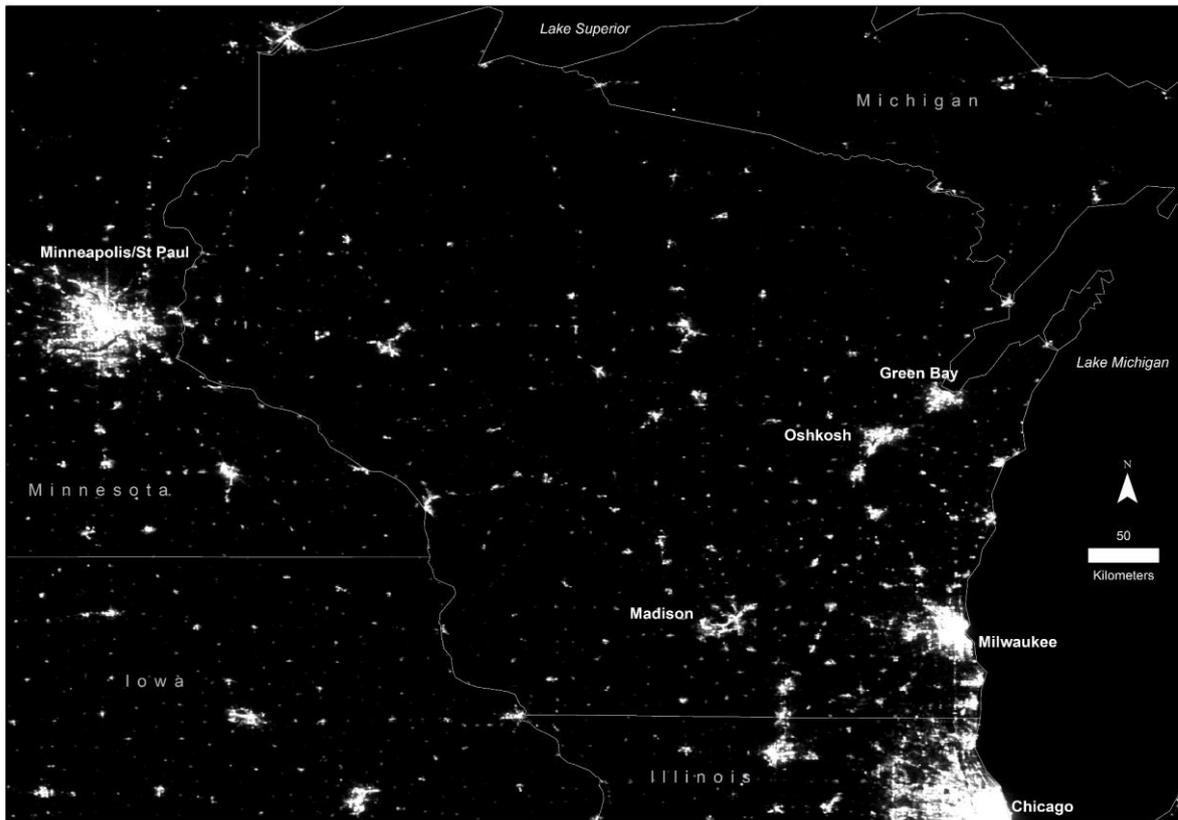
At least three studies estimate equation (1) with DMSP data as the noisy measure and the more accurate VNL or BM data as the truer values.⁴ With pixel-level data for urban areas (mostly in developing countries), Alimi et al (2022) estimate $\hat{\lambda} = 0.3$. A higher estimate, of $\hat{\lambda} = 0.7$, comes from regional data for Europe at the NUTS2 level (Gibson, 2021) which is typically the provincial level (or groups of counties in some countries). Kim et al (2022) use the second sub-national level for North Korea (which is counties in rural areas and districts in urban areas) and get $\hat{\lambda} = 0.4$. Although three studies is not much basis for extrapolation, one explanation for variation in $\hat{\lambda}$ is that mean reversion shows up more (as a lower $\hat{\lambda}$) for smaller spatial units because the blurring has a proportionally larger effect on small spatial units.

III. Visual Evidence from Wisconsin

The mean-reverting errors in the DMSP data are illustrated in Figure 1, by comparing with VNL images for Wisconsin in the United States. Wisconsin has a medium level of spatial inequality, ranking 21st of the 50 states for the Gini index of county GDP. It is useful for this illustration because it is surrounded on two sides by water and by the sparsely populated upper peninsula of Michigan, minimizing visual distraction due to light from elsewhere (the other lights shown are mainly from Chicago and Minneapolis/St Paul).

⁴ In support of the assumption that VNL or BM data provide truer values, Vuong (1989) non-nested tests based on the Kullback-Liebler information criterion show that models using VNL or BM data provide results that are closer to the truth than are models using DMSP data (Gibson, 2021; Zhang and Gibson, 2022).

Figure 1: Wisconsin at Night (Annual Composites for 2013)
A: VIIRS Night Lights



B. DMSP



Much of the population and economic activity in Wisconsin is concentrated in the lower southeast corner of the state. Specifically, if one draws a straight line from Green Bay through Oshkosh and Madison down to the state border with Illinois, there are 22 counties (of 72 in the whole state) that are either intersected by that line or are southeast of it. While those 22 counties only cover 22% of the state area, they contributed 74% of state GDP (as of 2013), and were home to 69% of state population (as of the 2010 census). When VNL data shown in panel A of Figure 1 are used, the share of total state luminosity contributed by the counties in the southeast corner is similar to the GDP and population shares, with 66% of the state total of VNL radiance coming from these 22 counties.

Unlike the clear definition of urban areas shown by VNL data in panel A of Figure 1, the image using DMSP data in panel B is very blurred, with large areas that are unlit in the VNL image appearing illuminated according to DMSP. By attributing lights to unlit places, DMSP data revert towards the mean because unlit and poorly lit areas have their luminosity overstated. The other feature of DMSP that causes mean-reversion is attenuated differences at the top of the lighting distribution, so that more brightly lit big cities seem no brighter than far smaller towns. For example, the brightness of Oshkosh (population 67,000) seems hardly less than that of Milwaukee, whose population is ten times larger, when the DMSP data are used (the maximum available DMSP digital number of 63 is recorded for both cities). Yet VNL data that are not subject to top-coding show the most luminous parts of Milwaukee have radiance at least twice that of Oshkosh (at 240 nano Watts per square centimetre per steradian ($\text{nW}/\text{cm}^2/\text{sr}$) in Milwaukee versus 110 $\text{nW}/\text{cm}^2/\text{sr}$ for Oshkosh).

As a consequence of these mean-reverting errors in the DMSP data, spatial inequality is greatly understated. Whereas GDP, population, and VNL data all show that at least two thirds of Wisconsin totals come from the 22 counties in the southeast corner of the state, the DMSP data suggest these counties contribute a minority of total luminosity, with 47% from the southeast corner and 53% from the rest of the state. By raising the share of lights that seem to come from the less economically active part of the state, and lowering the share that seem to come from the part with concentrated economic activity, the DMSP data create a systematic understatement of spatial inequality. Consequently, the Gini index for Wisconsin based on county-level data is only 0.27 according to DMSP while it is 0.51 if the VNL data are used (and slightly higher, at 0.57 when using GDP data).

IV. Inattention to Spatially Mean-Reverting Errors in the Literature

Despite the clear distortion seen in Figure 1, which also holds elsewhere (e.g. Gibson et al, 2020; Gibson, 2021) the inequality literature using DMSP data ignores the implications of spatially mean-reverting errors for inequality estimates. We reviewed 26 studies (15 of which had published since 2018) that use night-time lights data to measure spatial inequality. Amongst these studies, 22 used DMSP data. There is no decline over time in the popularity of this data source, as 12 of the 15 most recent studies used DMSP data. The full details on the studies reviewed, including the data source(s), time scale and spatial units, and research objective, are provided in Appendix A, with summary tabulation of the review in Table 2.

Table 2: Summary of the Studies Reviewed

	All studies	Recent studies (2018 onwards)	Earlier studies (prior to 2018)
Number of studies	26	15	11
Number that use DMSP data	22	12	10
Number that mention mean-reverting errors in DMSP data	0	0	0
Number that de-blur DMSP data	0	0	0
Number that adjust for top-coding in DMSP data	1	1	0

Note: Details on the studies summarized in this table can be found in Appendix A.

We carefully reviewed each study to see if authors mention spatially mean-reverting errors in the DMSP data. We also checked if any method was used to de-blur DMSP data, such as the procedures outlined in Abrahams et al (2018). Similarly, we checked if anything was done to correct the data for top-coding problems.⁵

There were no mentions of spatially mean-reverting errors in the studies we reviewed. Likewise, none used procedures to de-blur the DMSP data (Table 2). In terms of top-coding, one study (Montalvo et al, 2021) adjusted top-coded observations relative to the distribution of observations in the locality that were not top-coded. Two other studies excluded from their sample certain places that the authors assumed might be affected by top-coding; Chakravarty and Dehejia (2017) exclude five major metro areas (Mumbai, Kolkata, Chennai, Bangalore and Hyderabad) as likely having top-coded data, and the same reason is used by Mendez and Santos-Marquez (2020) for excluding Singapore. Such exclusions are likely to create biases due to sample selectivity and it is also the case that a study of regional inequality that excludes very densely populated metro areas may not be very informative for policy-makers. Moreover, the top-coding issues extend far further down the distribution of DMSP digital numbers, due to the fact that various pixels may be top-coded on some nights of the year but not on others (given the variation in sensor amplification over the lunar cycle) and so annual composites may have pixels that are partially top-coded even for annual average DN values as low as DN=55 (Bluhm and Krause, 2022).

Another symptom of the inequality literature ignoring mean-reverting errors in DMSP data is that some of the studies reviewed used a time-series that spliced together DMSP and VNL data (Chen and Zhang, 2022; Ivan et al, 2020). This splicing of the two data sources was to extend the length of their time-series past the 2013 ending data of the most popular DMSP annual composites.⁶ However, with the DMSP data having a mean-reverting error there is no

⁵ Bluhm and Krause (2022) released Pareto-adjusted DMSP data for 1992-2013 for all pixels worldwide which are meant to deal with top-coding (under the assumption that the true luminosity from the top-coded pixels follows a Pareto distribution). Those corrected data were available since the release of their first working paper in 2018. Likewise, a MATLAB script for the Abrahams et al deblurring approach was available since 2015.

⁶ An extended set of DMSP annual composites carrying on past 2013 has been provided by Ghosh et al (2021) but these data do not appear to have been used by any of the inequality studies. Unlike the original DMSP data

simple adjustment factor to ‘line up’ the two data sources (notwithstanding that some studies in the remote sensing literature attempt to create just such a concordance). The reason is that a mean-reverting measurement error is not constant and instead will vary with the true, but unknown, level of luminosity (as seen from equation (1), noting that the error, v varies with the true value). A lack of discussion of this issue by these inequality studies using spliced datasets is more evidence that implications of mean-reverting measurement errors in DMSP data seem to have been ignored.

V. County-level Evidence from the United States and China

The example given in Section III suggests that inequality estimates from DMSP data may be subject to a considerable downward bias. The discussion in Section II shows that the mean-reverting measurement errors that cause this downward bias in inequality estimates are due to inherent features of the DMSP sensors and data management, and so there is no reason to expect that the bias would not occur more widely beyond the particular example of Wisconsin used in Section III. Therefore, in this section we demonstrate that the bias is more generalized, using data from all parts of the United States and also from China, which are two of the few countries that report GDP at the county-level. The advantage of working with data from countries that have county-level GDP data is that it gives a benchmark for the inequality estimates derived from night-time lights data; if these county-level GDP data were available more widely they would surely be used for spatial inequality studies. The details on the GDP data sources we use for the two countries, and on the sources of the night-time lights data that we use, are provided in Appendix B.

Spatial inequality in the United States is shown in Figure 2 to have been quite stable over the last two decades according to either the Gini coefficient (panel A) or the Theil index (panel B). Specifically, the county-level GDP data provide Gini coefficients that hardly vary, ranging from a low of 0.71 in 2009 to a high of 0.72 in 2017 (with a mean of 0.713 across the 19 years with GDP data and a standard deviation across the annual estimates of 0.003). Along the same lines, annual values of the Theil index hardly varied, ranging from 0.98 to 1.01 (and the years with the minimum and maximum values are the same as for the Gini coefficient). The bootstrapped standard errors surrounding the inequality estimates for each year average 0.03 for the Gini coefficient and 0.12 for the Theil index.

In contrast to what is shown with the GDP data, the DMSP data on night-time lights suggest a significantly lower level of spatial inequality. The Gini coefficients vary from a high of 0.52 in 2006 to a low of 0.48 in 2010 (the mean is 0.50 and the standard deviation of the annual estimates is 0.011). There is no overlap between 95% confidence intervals of the DMSP-derived Gini coefficients and the GDP-based Gini coefficients. The gap between the lights-based and GDP-based estimates is even more apparent for the Theil index; the mean of this index is 0.44 when DMSP is used, which is almost 60% below what the GDP data show. Moreover, the average gap between the GDP-based Theil index and the DMSP-based index is

provided from 1992-2013, which were based on early evening observations, the extended series uses pre-dawn readings and so the composition of lights is likely to be different and this may affect apparent inequality.

over five times larger than the bootstrapped standard errors (for the Gini the average gap is six times larger than the standard errors). In other words, based on the patterns of economic activity that prevail across all counties in the United States, inequality measures derived from DMSP data are statistically significantly lower than the estimates derived from GDP data.

This gap in inequality estimates can be most correctly described as understatement when using DMSP data. The first reason for this claim is that GDP data would likely be the first choice of researchers wanting an economic activity measure, if such data were reliably available at a spatially disaggregated level elsewhere. In other words, the estimates derived from the GDP data can be sensibly thought of as a benchmark. Second, inequality estimates based on the VNL data show no significant differences from those based on GDP data.⁷ The discussion in Section II on DMSP sensors and data management gave good reasons to expect that the DMSP data will be less accurate than VNL and BM data and so close correspondence between the GDP-based and VNL-based estimates lends credence to the GDP-based results.

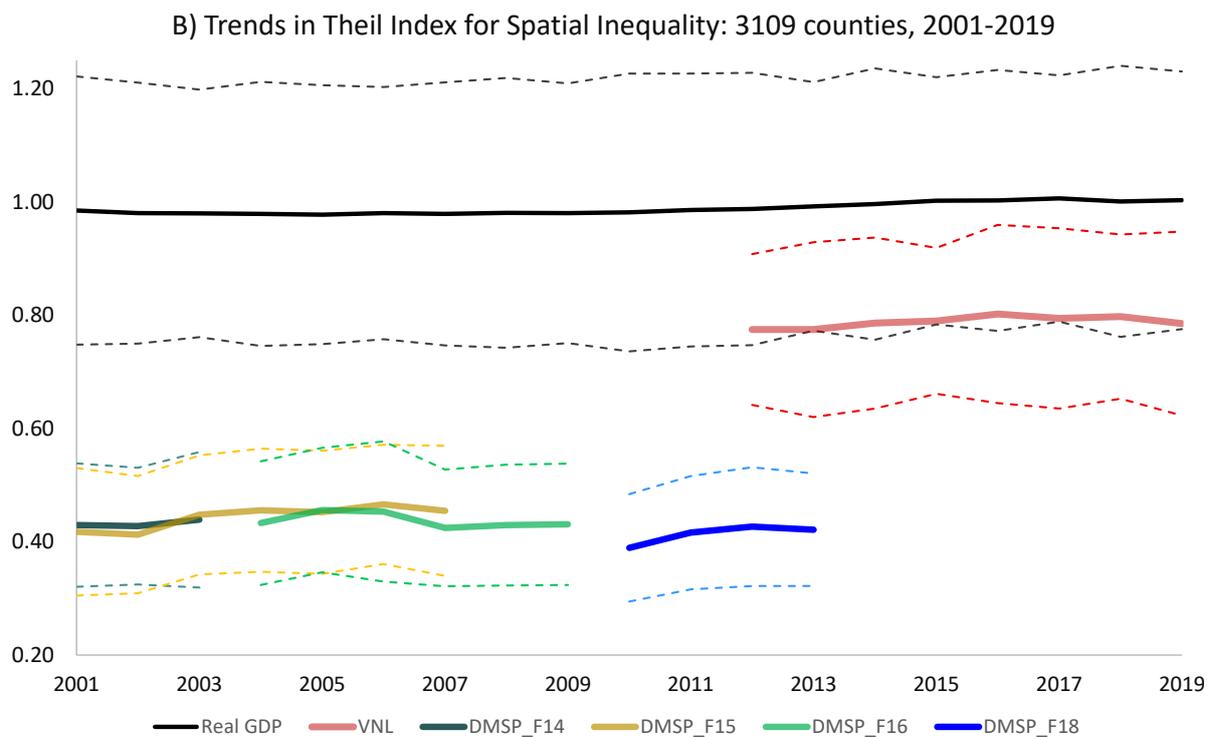
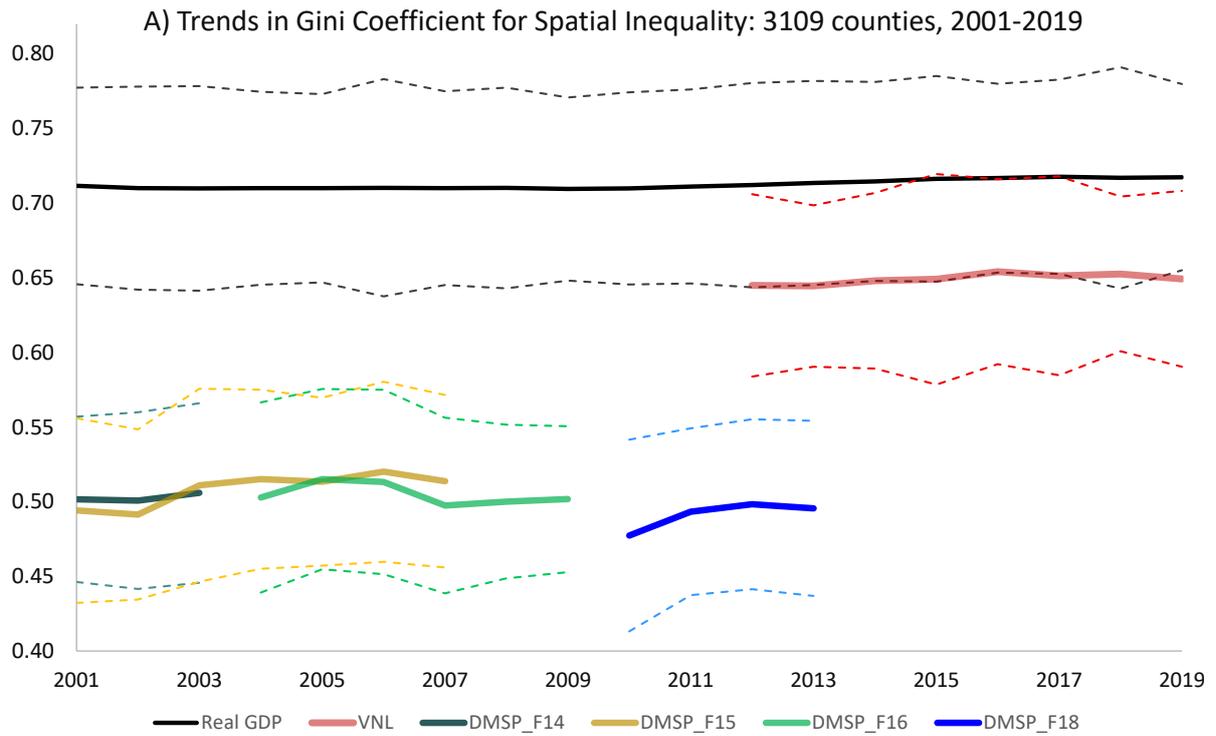
Two other features of Figure 2 warrant comment. First, the time-series of inequality estimates based on DMSP data has discontinuities in the years when DMSP transitions from one satellite to another. For example, when the source satellite switched from F16 in 2009 to F18 in 2010 there was a four point drop in the Theil index—equivalent to a ten percent decline in the apparent level of inequality—even though GDP-derived inequality estimates showed no similar change in those years. These discontinuities are likely to stem from a lack of calibration of the DMSP satellites which interferes with the consistency of the DMSP time-series (Gibson et al, 2020).⁸ This flaw also shows up as discrepancies when two satellites provide the data in the same year (e.g. spatial inequality in 2007 was seven percent lower with the data from satellite F16 than with what satellite F15 indicated for the same year).

The other feature of the Figure 2 results is that the understatement of inequality is far more apparent for the Theil index than for the Gini coefficient; probably because the Theil index is more sensitive to differences at the top of the distribution than is the Gini coefficient. There are at least two possible reasons why spatial variation at the top of the distribution is especially understated with the DMSP data; non-linearity in the GDP-luminosity relationship whereby DMSP night-time lights respond less to GDP in places where GDP is higher (Bluhm and McCord, 2022) and top-coding bias whereby DMSP data cannot distinguish brightly lit economic features, such as central business districts, from their less brightly lit surrounds (Bluhm and Krause, 2022; Gibson et al, 2021). A lack of uniform understatement of spatial inequality may distort the understanding of what drives some spatial inequality, given that it is not simply a parallel shift down in all inequality measures when the DMSP data are used.

⁷ We do not show the inequality estimates from BM data in Figure 2, in order to avoid clutter, but these are very similar to the inequality estimates from the VNL data, and quite distinct to those from the DMSP data.

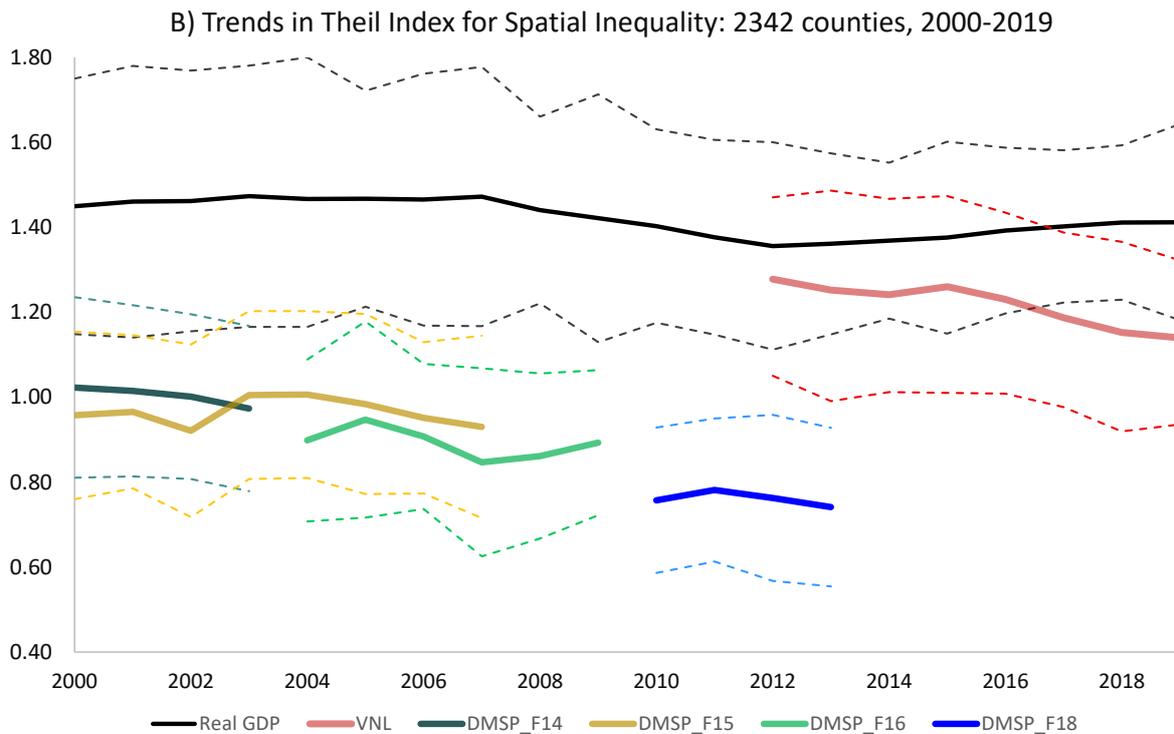
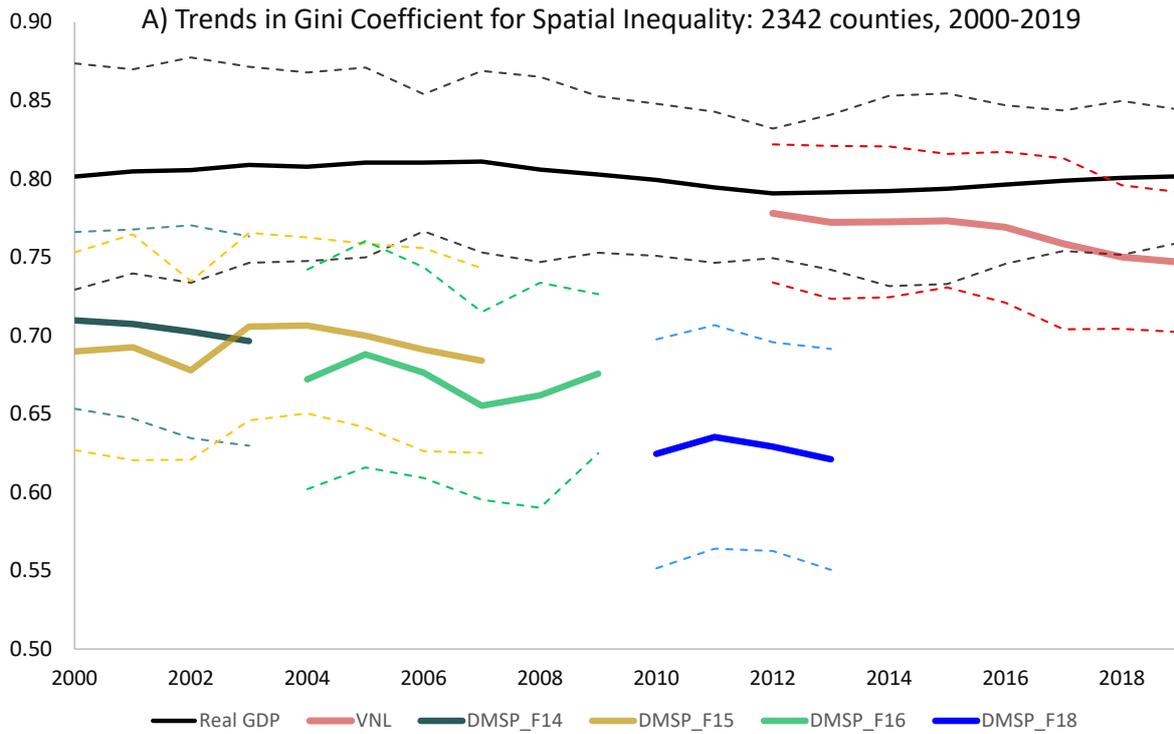
⁸ Remote sensing researchers often use ‘inter-calibrated’ DMSP data but economists ignore these data in favour of using year fixed effects (Gibson et al, 2020). The inter-calibration uses a regression to line-up the results from the various satellites. This just alters the first moment of the distribution so has no effect on inequality estimates.

Figure 2: County-level Spatial Inequality in the United States



Note: 95% confidence intervals from bootstrapped standard errors shown with dashed lines.

Figure 3: County-level Spatial Inequality in China



Note: 95% confidence intervals from bootstrapped standard errors shown with dashed lines.

The county-level inequality results for China are shown in Figure 3, and the patterns seen in the data for the United States are also exhibited for China. The VNL-based inequality estimates are not statistically significantly different to the GDP-based estimates; confidence intervals for the Gini coefficients with these two series almost completely overlap from 2012 to 2015, and overall the VNL-based Gini coefficients are just 4% (3 Gini points) below the GDP-based ones. In contrast, the DMSP-based inequality measures are 16% (37%) below the benchmark GDP-based estimates for the Gini coefficient (Theil index) over the 2000-2013 period that has both sets of data. The greater proportionate understatement of the Theil index (compared with the understatement of the Gini coefficient) if the DMSP data are used is in line with the pattern seen in Figure 2 for the United States.

In addition to the DMSP data indicating that the level of China's spatial inequality is lower than what either the GDP data or the more accurate VIIRS night-time lights data show, analyses of temporal trends in inequality could also be skewed by using the DMSP data. Over the 2000-2013 period there was a slight fall in inequality according to GDP data—the Gini coefficient went from 0.80 to 0.79 and the Theil index from 1.45 to 1.36—but far larger falls in inequality for that period are implied by the DMSP data. The DMSP-derived Theil index went from 1.02 in 2000 to 0.74 in 2013, which was a 28% fall (compared to a fall of just 6% in the GDP-based measure for the same period) Likewise the DMSP-derived Gini coefficient went from 0.71 to 0.62, and the trend annual rate of decline in the Gini coefficient based on DMSP data for the 2000-2013 period was -1% (with standard error of -0.1%). This annual trend rate of decline with the DMSP data is statistically significantly ($p < 0.01$) greater (i.e., it implies an apparently faster decline) than is the annual trend in the inequality estimates that are derived from the GDP data for China.⁹ In other words, the impact that the errors in the DMSP data have on inequality estimates cannot just be treated as an intercept shift that affects all years (and all inequality measures) equally. Instead, these measurement errors in DMSP data also seem to affect evidence on the temporal trends in spatial inequality.

VI. Conclusions

Researchers in economics and regional science increasingly use satellite-detected night-time lights data; mostly from the Defense Meteorological Satellite Program whose original purpose was cloud observation for short-term weather forecasting. Studies using these data are part of a growing 'big data' movement in social science research. Yet despite the popularity of these DMSP data they have inherent measurement error features whose implications are largely ignored. In terms of measuring on-the-ground economic activity the DMSP data are flawed by blurring and top-coding, due to intrinsic limitations of the sensors and data management. While prior authors have noted that even though DMSP data may be a noisy proxy they can nevertheless still be useful (e.g. Henderson et al, 2012), the fact that the measurement errors

⁹ The United States also showed this pattern although it is less visually apparent in Figure 2. With DMSP data the Gini coefficient was declining at a trend annual rate of -0.2% and the Theil index at -0.4% (with both trends statistically significant at $p < 0.05$) over the 2001-13 period. These declines were both statistically significantly faster (at the $p < 0.04$ level) than the trends in the GDP-based inequality measures. Running the same test for differences in time trends with the VNL-based inequality measures from 2012-19 showed no difference in the VNL-based trend compared with what the GDP-based inequality measures showed over that period.

are spatially mean-reverting, and the implications of this error pattern for particular uses of these data, are not emphasized. In particular, spatially mean-reverting errors make the DMSP data singularly ill-suited for estimating spatial inequality, even if they may adequately proxy for the average levels of (or growth in) economic activity at the national or aggregated regional level in other research contexts.

In this paper we explain the sources of these mean-reverting errors and give examples of their impacts on spatial inequality estimates for two countries that have county-level GDP data available to use as a benchmark. For the United States we find that the Theil index of inequality is understated by almost 60% if the DMSP data are used, compared to what GDP data show. The understatement of the Gini coefficient is somewhat less but still statistically significant. For China, the Theil index appears to have fallen by 28% over the 2000 to 2013 period if the DMSP data are used—where this fall implies some levelling of China’s spatial distribution of economic activity at a time when various government initiatives aimed to promote lagging regions. However, the GDP data do not show a similarly rapid reduction in spatial inequality and so the possibility that policy makers could be inadvertently misled by studies that use DMSP data is quite strong.

In addition to these examples, which provide some grounds for pausing use of DMSP data for estimating spatial inequality statistics, we reviewed existing studies to see whether these measurement error issues have been considered by inequality researchers. Our review showed that the literature to date appears to have ignored these issues, with no discussion of the implications of spatially mean-reverting errors in DMSP data for inequality estimates. In addition, deblurring approaches for DMSP data available since 2015 (Abrahams et al, 2018) and data adjusted for top-coding (for all DMSP pixels from 1992 to 2013) that are available since 2018 (Bluhm and Krause, 2022) have not been used, despite many of the studies that we reviewed post-dating the availability of these correction methods. The fact that inequality studies have continued to use DMSP data when the far more accurate VIIRS Night Lights and Black Marble data are available also suggests that the measurement error problems with DMSP data have not been given sufficient attention. Thus, even though the availability of satellite-detected night-time lights data has the potential to overcome some data constraints that historically have limited the range of spatial inequality studies that can be carried out, the uncritical use of DMSP data may introduce some new misunderstandings about the levels of, and trends in, spatial inequality. It is especially in settings where conventional indicators of economic activity, such as GDP and household survey data, are unavailable or untrustworthy that the measurement errors in the DMSP data may contribute to distorted regional policies because the paucity of data in such places limits the opportunity to re-examine findings that rely on DMSP data.

References

- Abrahams, Alexei, Christopher Oram and Nancy Lozano-Gracia (2018) Deblurring DMSP night-time lights: A new method using Gaussian filters and frequencies of illumination. *Remote Sensing of Environment* 210(1): 242-258.
- Alimi, Omoniyi, Geua Boe-Gibson, and John Gibson (2022) Noisy night lights data: Effects on research findings for developing countries. *Working Paper 12/22*, Department of Economics, University of Waikato.
- Black, Dan, Mark Berger, and Frank Scott (2000) Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association* 95(451): 739-748.
- Bluhm, Richard, and Gordon McCord (2022) What can we learn from nighttime lights for small geographies? Measurement errors and heterogeneous elasticities. *Remote Sensing* 14(5): 1190.
- Bluhm, Richard, and Melanie Krause (2022) Top lights: Bright cities and their contribution to economic development. *Journal of Development Economics* 157(1): 102880.
- Chakravarty, Praveen, and Vivek Dehejia (2017) India's income divergence: Governance or development model? *Briefing Paper Series*, IDFC Institute.
- Chen, Guohui, and Jie Zhang (2022) Regional inequality in ASEAN countries: Evidence from an outer space perspective. *Emerging Markets Finance and Trade* (online early available at: <https://doi.org/10.1080/1540496X.2022.2119810>)
- Deininger, Klaus, and Lyn Squire (1996) A new data set measuring income inequality. *The World Bank Economic Review* 10(3): 565-591.
- Doll, Christopher (2008) *CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications*. Center for International Earth Science Information Network, Columbia University, New York.
- Elbers, Chris, Jean O. Lanjouw, and Peter Lanjouw (2003) Micro-level estimation of poverty and inequality. *Econometrica* 71(1): 355-364.
- Elvidge, Christopher, Kimberly Baugh, Mikhail Zhizhin, and Feng-Chi Hsu (2013) Why VIIRS data are superior to DMSP for mapping night time lights. *Proceedings of the Asia-Pacific Advanced Network* 35(1): 62-69.
- Elvidge, Christopher, Mikhail Zhizhin, Tilottama Ghosh, Feng-Chi Hsu, and Jay Taneja (2021) Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sensing* 13(5): 922.
- Ghosh, Tilottama, Kimberly Baugh, Christopher Elvidge, Mikhail Zhizhin, Alexey Poyda, and Feng-Chi Hsu (2021) Extending the DMSP nighttime lights time series beyond 2013. *Remote Sensing* 13(24): 5004.
- Gibson, John (2021) Better night lights data, for longer. *Oxford Bulletin of Economics and Statistics* 83(3): 770-791.
- Gibson, John, Jikun Huang, and Scott Rozelle (2001) Why is income inequality so low in China compared to other countries? The effect of household survey methods. *Economics Letters* 71(3): 329-333.
- Gibson, John and Bonggeun Kim (2010) Non-classical measurement error in long-term retrospective recall surveys. *Oxford Bulletin of Economics and Statistics* 72(5): 687-695.

- Gibson, John, Susan Olivia, and Geua Boe-Gibson (2020) Night lights in economics: Sources and uses. *Journal of Economic Surveys* 34(5): 955-980.
- Gibson, John, Susan Olivia, Geua Boe-Gibson and Chao Li (2021) Which night lights data should we use in economics, and where? *Journal of Development Economics* 149(1): 102602.
- Henderson, Vernon, Adam Storeygard, and David Weil (2012) Measuring economic growth from outer space. *American Economic Review* 102(2): 994-1028.
- Hsu, Feng-Chi, Kimberly Baugh, Tilottama Ghosh, Mikhail Zhizhin, and Christopher Elvidge (2015) DMSP-OLS radiance calibrated night-time lights time series with inter-calibration. *Remote Sensing* 7(2): 1855-1876.
- Ivan, Kinga, Iulian-Horia Holobacă, József Benedek, and Ibolya Török (2020) Potential of night-time lights to measure regional inequality. *Remote Sensing* 12(1): 33.
- Kim, Bonggeun, John Gibson and Geua Boe-Gibson (2022) How effective are sanctions on North Korea? Popular DMSP night-lights data may bias evaluations due to blurring and poor low-light detection. *Working Paper Working Paper* 06/22, Department of Economics, University of Waikato.
- Mendez, Carlos, and Felipe Santos-Marquez (2020) Regional convergence and spatial dependence across subnational regions of ASEAN: Evidence from satellite nighttime light data. *Regional Science Policy & Practice*, 13(6): 1750-1777.
- Michalopoulos Stelios, and Elias Papaioannou (2014) National institutions and sub-national development in Africa. *Quarterly Journal of Economics* 129(1): 151-213.
- Montalvo, José, Marta Reynal-Querol, and Juan Carlos Muñoz Mora (2021) Measuring inequality from above. *Working Papers* 1252, Barcelona School of Economics.
- Román, Miguel, Zhuosen Wang, Qingsong Sun, Virginia Kalb, Steven Miller, Andrew Molthan, ..., and Edward Masuoka (2018). NASA's Black Marble nighttime lights product suite. *Remote Sensing of Environment* 210 (1): 113-143.
- Sangkasem, Krittaya, and Nattapong Puttanapong (2022). Analysis of spatial inequality using DMSP-OLS nighttime-light satellite imageries: A case study of Thailand. *Regional Science Policy & Practice* 14(4): 828-849.
- Sharp, Michael, Bertrand Buffière, Kristen Himelein, Nathalie Troubat, and John Gibson (2022) Effects of data collection methods on estimated household consumption and survey costs. *Policy Research Working Paper* No. 10029 (World Bank, Washington).
- Small, Christopher, Francesca Pozzi, and Christopher D. Elvidge (2005) Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sensing of Environment* 96(3-4): 277-291.
- Tuttle, Benjamin, Sharolyn Anderson, Paul Sutton, Christopher Elvidge, and Kim Baugh (2013) It used to be dark here. *Photogrammetric Engineering & Remote Sensing* 79(3): 287-297.
- Vuong, Quang (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2): 307-333.
- Zhang, Xiaoxuan, and John Gibson (2022) Using multi-source nighttime lights data to proxy for county-level economic activity in China from 2012 to 2019. *Remote Sensing* 14(5): 1282.

Appendix A: Selected studies using satellite-detected night-time lights to measure spatial inequality (most recent first)

Authors	Year	Data source	Spatial units	Time period	Deblur?	Adjust for top-coding?	Mention mean-reverting errors?	Research Objective
Chen & Zhang	2022	DMSP(2000-13); VIIRS(2014-18)	337 regions in ten ASEAN countries	2000-2018	No	No	No	Investigate trends in subnational inequality in ASEAN countries.
Sangkasem & Puttanapong	2022	DMSP	76 provinces in Thailand	1994-2013	No	No	No	Compare inequality based on night-time lights with survey-based index.
Galimberti et al.	2021	DMSP	234 countries	1992-2013	No	No	No	Study economic inequality drivers and dynamics
Mirza et al.	2021	DMSP	Global and 50 States in USA	Five-yearly from 1990	No	No	No	Test if remotely sensed luminosity reflects economic inequality.
Montalvo et al.	2021	DMSP	186 countries	2000-2013	No	Yes	No	Develop new measure of night-time lights-based inequality.
Weidmann & Theunissen	2021	VIIRS	37 countries in Africa.	2012-2019	No	No	No	Compare inequality based on night-time lights with survey estimates.
Ivan et al.	2020	DMSP(1992-2013); VIIRS(2014-18)	42 administrative counties in Romania	1992-2018	No	No	No	To measure regional inequality in Romania.
Mendez-Guerra & Santos-Marquez	2020	DMSP	274 subnational regions in ASEAN	1998-2012	No	No	No	Evaluate subnational convergence and regional inequality.
Singhal et al.	2020	DMSP	Districts in 23 states in India	1999-2008	No	No	No	Test if regional inequality based on night-time lights has Kuznets Curve.
Elvekjær	2019	VIIRS	402 regions in 36 OECD countries	2015	No	No	No	Compare models that predict GDP & inequality at five spatial scales.
Mukhopadhyay & Urzainqui	2018	DMSP	Villages, urban blocks & districts in India	2004 & 2011	No	No	No	Decompose inequality into within- and between-area components.
Mveyange	2018	DMSP	623 regions in 48 African countries	1992-2012	No	No	No	Small-area imputation: lights-based inequality used to predict income inequality where unreported.
Puttanapong & Zin	2018	DMSP	14 provinces in Myanmar	1992-2013	No	No	No	Estimate trends in regional inequality in Myanmar.
Sangkasem	2018	DMSP	76 provinces in Thailand	1992-2013	No	No	No	Reveal the clusters of poverty and inequality in Thailand.
Wu et al.	2018	VIIRS	341 prefectures & 31 provinces in China	2014-2017	No	No	No	Estimate trends in Theil index & in inequality components
Chakravarty & Dehejia	2017	DMSP	676 districts in India	1992-2013	No	No	No	Test for intra-state divergence in economic activity
Hu et al.	2017	DMSP	13 prefectural cities in Beijing-Tianjin-Hebei	2000-2012	No	No	No	Estimate trends in spatial inequality and identify spatial clusters.
Lessmann & Seidel	2017	DMSP	3163 regions in 180 countries	1992-2012	No	No	No	Study convergence and determinants of regional inequality.
Li et al.	2017	DMSP	41 regions in five Cent. Asian countries	1993-2012	No	No	No	Identify poor areas and measure regional development inequality.
Weidmann & Schutte	2017	DMSP	39 less developed countries	2003-2012	No	No	No	Test if luminosity can predict local wealth, as detected by surveys
Alesina et al.	2016	DMSP	Ethnic group areas in 173 countries	1992-2009	No	No	No	Examine linkages between ethnicity, inequality, & economic development
Mveyange	2015	DMSP	423 regions in 32 African countries	1995, 2000, 2005	No	No	No	Test if luminosity can be used to proxy for regional inequality.
Xu et al.	2015	DMSP	8 regions, 31 provinces, 354 prefectures in China	2005 and 2010	No	No	No	Measure regional differences in inequality of public services.
Zhou et al.	2015	VIIRS	30 provinces and municipalities in mainland China	2012	No	No	No	Study socioeconomic inequality patterns in China using luminosity data.
Gennaioli et al.	2014	DMSP	1528 regions from 83 countries	1992-2010	No	No	No	Investigate how within-country regional inequality drives regional convergence across countries.
Elvidge et al.	2012	DMSP	Global, national (189), subnational (4550), & gridded.	2006	No	No	No	Explore potential of nighttime lights data for estimating disparities in income distribution.

References for Appendix A (most recent first)

- Chen, Guohui, and Jie Zhang (2022) Regional inequality in ASEAN countries: Evidence from an outer space perspective. *Emerging Markets Finance and Trade* (online early available at: <https://doi.org/10.1080/1540496X.2022.2119810>)
- Sangkasem, Krittaya, and Nattapong Puttanapong (2022) Analysis of spatial inequality using DMSP-OLS nighttime-light satellite imageries: A case study of Thailand. *Regional Science Policy & Practice* 14(4): 828-849.
- Galimberti, Jaqueson, Stefan Pichler, and Regina Pleningner (2021) Measuring inequality using geospatial data. *KOF Working Papers 493*. KOF Swiss Economic Institute, ETH Zurich.
- Mirza, M. Usman, Chi Xu, Bas van Bavel, and Marten Scheffer (2021) Global inequality remotely sensed. *Proceedings of the National Academy of Sciences* 118(18): e1919913118.
- Montalvo, José, Marta Reynal-Querol, and Juan Carlos Muñoz Mora (2021) Measuring inequality from above. *Working Papers* 1252, Barcelona School of Economics.
- Weidmann, Nils, and Gerlinde Theunissen (2021) Estimating local inequality from nighttime lights. *Remote Sensing* 13(22): 4624.
- Ivan, Kinga, Iulian-Horia Holobacă, József Benedek, and Ibolya Török (2020) Potential of night-time lights to measure regional inequality. *Remote Sensing* 12(1): 33.
- Mendez, Carlos, and Felipe Santos-Marquez (2020) Regional convergence and spatial dependence across subnational regions of ASEAN: Evidence from satellite nighttime light data. *Regional Science Policy & Practice*, 13(6): 1750-1777.
- Singhal, Abhishek, Sohini Sahu, Siddhartha Chattopadhyay, Abhijit Mukherjee, Soumendra Bhanja (2020) Using night time lights to find regional inequality in India and its relationship with economic development. *PLOS ONE*, 15(11): 1-18.
- Elvekjaer, Neja Maegaard (2019) Disaggregating economic inequality from space. *Thesis*, Department of Physical Geography and Ecosystem Science, Lund University.
- Mukhopadhyay, Abhiroop, and Garcés Urzainqui David (2018) The dynamics of spatial and local inequalities in India. *WIDER Working Paper*, United Nations University (UNU), World Institute for Development Economics Research (WIDER).
- Mveyange, Anthony (2018) Measuring and explaining patterns of spatial income inequality from outer space: Evidence from Africa. *World Bank Policy Research Working Paper* No. 8484, Development Economics Research Group, The World Bank.
- Puttanapong, Nattapong, and Shwe Zin Zin (2018) Spatial inequality in Myanmar during 1992-2016: An application of spatial statistics and satellite data. *The Social Science Review: Special Issue on Mekong Economy*, 156: 161-182.
- Sangkasem, Krittaya (2018) Poverty and inequality assessment using DMSP/OLS nighttime light satellite imageries at provincial level in Thailand. *Thesis*, Thammasat University.
- Wu, RongWei, Degang Yang, Jiefang Dong, Lu Zhang, and Fuqiang Xia (2018) Regional inequality in China based on NPP-VIIRS night-time light imagery. *Remote Sensing*, 10(2): 240.
- Chakravarty, Praveen, and Vivek Dehejia (2017) India's income divergence: Governance or development model? *Briefing Paper Series*, IDFC Institute.
- Hu, Yi'na, Jian Peng, Yanxu Liu, Yueyue Du, Huilei Li, and Jiansheng Wu (2017) Mapping

- development pattern in Beijing-Tianjin-Hebei urban agglomeration using DMSP/OLS nighttime light data. *Remote Sensing*, 9(7): 760.
- Lessmann, Christian, and Seidel André (2017) Regional inequality, convergence, and its determinants—A view from outer space. *European Economic Review*, 92: 110-132.
- Li, Shuang, Tongyao Zhang, Zhiyu Yang, Xi Li, and Huimin Xu (2017) Night time light satellite data for evaluating the socioeconomics in Central Asia. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42: 1237-1243.
- Weidmann, Nils B, Sebastian Schutte (2017) Using night light emissions for the prediction of local wealth. *Journal of Peace Research*, 54(2): 125-140.
- Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou (2016) Ethnic inequality. *Journal of Political Economy*, 124(2): 428-488.
- Mveyange, Anthony (2015) Night lights and regional income inequality in Africa. *WIDER Working Paper*, The United Nations University World Institute for Development Economics Research (UNU-WIDER), Helsinki.
- Xu, Huimin, Hutao Yang, Xi Li, Huiran Jin, and Deren Li (2015) Multi-scale measurement of regional inequality in Mainland China during 2005–2010 using DMSP/OLS night light imagery and population density grid data. *Sustainability*, 7(10): 13469-13499.
- Zhou, Yuke, Ting Ma, Chenghu Zhou, and Tao Xu (2015) Nighttime light derived assessment of regional inequality of socioeconomic development in China. *Remote Sensing*, 7(2): 1242-1262.
- Gennaioli, Nicola, Rafael La Porta, Florencio Lopez De Silanes, and Andrei Shleifer (2014) Growth in regions. *Journal of Economic Growth*, 19(3): 259-309.
- Elvidge, Christopher, Kimberly Baugh, Sharolyn Anderson, Paul Sutton, and Tilottama Ghosh (2012) The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data. *Social Geography*, 7(1): 23-35.

Appendix B: Details on the Data Sources for the County-level Analysis in Section V

The county-level GDP data for the United States are from the Bureau of Economic Analysis (BEA) at: <https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas>. We used chained series in 2012 dollars. Annual estimates are provided separately for each county for the 2001 to 2019 period, with three broad exceptions. In Alaska the BEA combine some census areas in their reporting, and in Hawaii they combine Maui and Kalawao counties. Virginia has the most adjustments because the BEA create 23 combination areas where one or two independent cities whose population in the 1980 census was less than 100,000 are combined with an adjacent county. Overall there are $n=3109$ counties and combination areas (we refer to all of these as county-level units or just counties) with data available in each year.

In contrast to the United States, sub-national GDP data for China are not provided by a single unified source. Instead we used three types of publications from the National Bureau of Statistics (NBS) to build our GDP database: annual editions of the China Statistical Yearbook (county-level) (in Chinese it is *Zhongguo Xianyu Tongji Nianjian (Xianshi Juan)*), annual editions of the China City Statistical Yearbook (known in Chinese as *Zhongguo Chengshi Tongji Nianjian*), and annual editions of the Statistical Yearbook for each city or province (for example the Beijing Statistical Yearbook) (NBS, various dates). Each edition reports on GDP the previous year, so we use the 2001 to 2020 editions to obtain annual GDP data from 2000 to 2019. When any county (the typical spatial unit in rural areas) or district (the typical unit in urban areas) was subsequently merged we enforce the same aggregation on earlier years to have a consistent 2000-19 time-series for each spatial unit. Overall, we have annual GDP data for each of $n=2342$ units at the 3rd level of the sub-national administrative hierarchy, where these units maintain a consistent spatial definition from 2000 to 2019. We refer to these as either county-level units or just counties.

The DMSP data are annual composites from satellites F14, F15, F16 and F18 that collectively cover each year from 2000 to 2013 (with some years having two satellites providing data). The stable lights product (where ‘stable’ simply means ephemeral lights such as from fires are removed, it does not necessarily imply temporal consistency) provides 6-bit digital numbers (DN) ranging from 0–63. The technical details on the steps used to create these data are in Baugh et al (2010) and the data are available at: <https://eogdata.mines.edu/products/dmsp/>.

The VNL data are version 2.1 annual composites for 2012 to 2019 from Elvidge et al (2021) based on monthly cloud-free radiance averages coming from the Suomi/NPP satellite. These data undergo an initial filtering to remove extraneous features such as fires and aurora before the resulting rough annual composites have further outlier removal procedures applied. The lit grid cells are isolated from background noise using multi-year thresholds that make these data better for change detection than earlier vintages of VNL data. The data are in units of nano Watts per square centimeter per steradian ($nW/cm^2/sr$) presented on a 15 arc-second output grid available from: https://eogdata.mines.edu/nighttime_light/annual/v21/.

The NASA Black Marble annual composites are derived from the same satellite as VNL V2.1 data but are processed differently (Román et al, 2018). The data products are corrected for atmospheric, terrain, vegetation, snow, lunar, and stray light effects on the radiance values, which are calibrated across time and are also validated against ground measurements. The data are in units of nano Watts per square centimeter per steradian ($nW/cm^2/sr$), with 16-bit precision on a 15 arc-second output grid. We use the all-angle composites for snow-free nights, available from: <https://ladsweb.modaps.eosdis.nasa.gov/archive/allData/5000/VNP46A4/>.

References

- Baugh, K., Elvidge, C., Ghosh, T., & Ziskin, D. (2010). Development of a 2009 stable lights product using DMSP-OLS data. *Proceedings Asia-Pacific Advanced Networks* 30, 114.
- Elvidge, C., Zhizhin, M., Ghosh, T., Hsu, F.-C., & Taneja, J. (2021). Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sensing* 13, 922.
- National Bureau of Statistics [NBS]. China Statistical Yearbook (county-level)[M]. Beijing: China Statistics Press, 2001-2020.
- National Bureau of Statistics [NBS]. China City Statistical Yearbook[M]. Beijing: China Statistics Press, 2001-2020.
- National Bureau of Statistics [NBS]. Beijing Statistical Yearbook[M]. Beijing: China Statistics Press, 2001-2020.
- Román, M., Wang, Z., Sun, Q., Kalb, V., Miller, S., Molthan, A., Schultz, L., Bell, J., Stokes, E., Pandey, B., & Seto, K. (2018). NASA's Black Marble nighttime lights product suite. *Remote Sensing of the Environment*, 210, 113-43.