Three Essays on Economic Applications Using Satellite Imageries of Nighttime Lights

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Leonardo José Maldonado Salazar

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Stephen Polasky, Advisor

December, 2023

Acknowledgements

Countless people have shaped my graduate experience, and this section is my humble attempt to express my heartfelt appreciation to as many of them as possible.

I am deeply grateful to my dissertation committee. Thank you to Stephen Polasky, my advisor, for his unwavering support, patience, guidance, and encouragement throughout my doctoral journey. I want to express special appreciation to Paul Glewwe for his insightful feedback and whose dedication to teaching inside and outside the classroom has been evident since day one. Thank you to Justin Johnson and Joe Knight for constructive comments and for sharing detailed references to deepen my knowledge of remote sensing methods and data-generating mechanisms. The careful and timely feedback of this committee significantly enhanced the quality of my work.

I also greatly benefited from the lessons and courses taught by Marc Bellemare, Jay Coggins, Elizabeth Davis, Frances Homans, Terrance Hurley, Larry Jones, Timothy Kehoe, Jason Kerwin, Gerard McCullough, Samuel Myers Jr., Hikaru Peterson, Rodney Smith, Chengyan Yue, as well as from discussions with colleagues and classmates in the Econometrics Frontier Group, Johnson-Polasky Lab Group, departmental and student seminars. They all had the role of contributing to my academic and research formation.

My heartfelt thanks to Aaron Sojourner from the Carlson School's Center for Human Resources and Labor Studies at the University of Minnesota and David Simon from the University of Connecticut. I had the honor of working and participating in technical discussions with them during my second year in the program. This opportunity complemented my tools to carry out research design and impact evaluation and allowed me to support my academic development financially. Special thanks also to colleagues at the Inter-American Development Bank, where I worked as a consultant in the last years of the program. They gave me constant support, stimulating discussions, and the flexibility

needed to complete this milestone.

I extend my warm thank you to the friends I made along the way, who have stood by me during this challenging yet rewarding phase of my academic life. Your encouragement and chatting fostered an intellectually engaging and less stressed environment. I would also like to thank old friends who recommended me and truly convinced me to follow a doctoral path: Ana Cuesta, Osmel Manzano, and José Pineda.

Finally, I am indebted to my family for all the emotional support and understanding during all these years. Your belief in me has been my driving force, and I am profoundly grateful for the sacrifices you have made to help me pursue my dreams.

Dedication

Para mis padres, Auristela Salazar de Maldonado y José Maldonado, para mi hermano, José Maldonado Salazar, y para mi esposa, Marbella Toro. Los llevo presente en mi mente y corazón en cada paso que doy.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In an era marked by technological improvement and advancements and a growing reliance on data-driven methodologies, the intersection of economics and remote sensing has emerged as a powerful component for comprehending and quantifying temporal and spatial dimensions of socioeconomic indicators. This dissertation aims to unravel applications in the economic field derived from analyzing nighttime light data. By delving into three chapters, I explore diverse ways satellite-recorded night light emissions (or lack thereof) can facilitate our understanding of different facets of economic phenomena.

Chapter 2, "Lighting-Up the Economic Activity of Oil-Producing Regions: A Remote Sensing Application," explores the relationship between nighttime lights and economic activity and analyzes whether there is a different relationship between oil-producing and non-oil producing regions. Oil regions may differ in how changes in the intensity of light relate to variations in economic activity because the oil industry has specific features that may strongly influence institutional, financial, and political subnational frameworks. I harmonized satellite datasets of night lights to construct a panel for eight oil-producing countries with regional data from 1992 to 2019. I find differences in the predictive power of lights by region and confirm how light data could improve economic growth measures. Even in an extreme scenario in which lights have no improvement in measuring true -with no measurement errors- growth in oil regions, they still improve estimates for non-oil areas.

In Chapter 3, "Living in Darkness: Rural Poverty in Venezuela," uses nighttime light imagery and gridded population datasets to estimate Venezuela's 2000–2020 rural poverty rates at the state and municipality levels. Then, I examine if there has been a significant change in rural poverty during the economic collapse ongoing since 2013–2014. The main finding reveals that most Venezuelan territory experienced a considerable increase in rural poverty rates between 2014 and 2020. Furthermore, I confirm how new rural poor areas appear across the country in clusters, surrounding municipalities with moderate to high poverty rates. This suggests that in recent years, more Venezuelans have sunk into darkness.

Finally, Chapter 4, "Measuring Regional Inequality in the Andean Countries: A Multiple-Stage Nested Theil Decomposition Using Night Light Emissions," examines inequality in the Andean countries using satellite-recorded nighttime lights and gridded population datasets from 2012 to 2021. It follows a multiple-stage nested Theil index

decomposition method accounting for each country's lowest administrative divisions to identify the appropriate local scale of policy actions for addressing inequality. The main findings reveal a decrease in overall inequality for the Andean region throughout the period primarily driven by a decline in between-country inequality and an increase in the relative importance of within-country inequality. In addition, there are spatial heterogeneities by country. Bolivia, Colombia, and Peru experienced a decline in wealth inequality over the past decade due to decreased disparities between provinces and less inequality within municipalities and districts, respectively. In contrast, the inequality components in Ecuador and Venezuela exhibit a more balanced contribution to overall inequality. And, while Ecuador does not show a significant change in overall inequality during the period, the inequality increase in Venezuela is primarily driven by changes in the disparity between all geographic subgroups. These findings provide insights for targeted initiatives to tackle inequality locally in the Andean region.

This dissertation seeks to contribute to the growing fields of economic and remote sensing through the lens of nighttime lights. By unraveling the complex interplay between economic activity, poverty, and inequality, these three chapters collectively offer a crucial exploration of potential applications of nighttime light imageries in understanding and addressing economic challenges. As the world evolves, this research aims to guide scholars and policymakers toward informed decision-making grounded in the luminosity of data-driven insights.

Chapter 2

Lighting-Up the Economic Activity of Oil-Producing Regions: A Remote Sensing Application

2.1 Introduction

The identification, measurement, and analysis of objects or areas from a distance might reveal characteristics not attainable otherwise. The use of remote sensors from a satellite, aircraft, or sonar systems allows us to record and to collect data from nature (mapping ocean features, monitoring temperature, clouds, storms, erosion) and to track the human impact on our surroundings (deforestation, mining, energy consumption). This tool is important for economics, where it is recognized the linkage between human settlements and economic growth. In this context, satellite-recorded nightlight data could be used as a proxy for economic activity. The general assumption is that nightlight emissions implicitly capture information about economic activity. The aggregate gross domestic product (GDP) is often badly measured, and the presence of measurement errors in official calculations requires alternative methodologies (Nordhaus and Chen 2012; Henderson et al. 2012).

This chapter explores the relationship between nighttime lights and economic activity, distinguishing between oil and non-oil producing regions. In particular, I calculate different elasticities between nighttime lights and economic activity in those regions. I examine if nighttime lights have different predictive power by region and if their use could improve the ability to measure true economic growth (i.e., a weighted optimal combination of observed and estimated growth from using light data). It is often assumed the presence of similar measurement errors in GDP data within each country, this chapter analyzes if that applies to oil-dependent economies or it is time to start accounting for subnational particularities.

In remote sensing literature, studies about oil-dependent countries are scarce. Recently, Do et al. (2017) and Debbich (2019) shed light on how oil activity could be measured using remote sensing methods. Do et al. (2017) use satellite multi-spectral imaging and ground-truth pre-war output data to construct a proxy for oil production in areas controlled by the Islamic State. Debbich (2019) looks to assess oil and non-oil GDP growth in Yemen between 2012 and 2017. As mentioned by Debbich (2019), studies based on nighttime light data make implicit assumptions on national income elasticities to different sources of light, including gas flaring. This situation shadows any analysis trying to distinguish specificities in economic growth.

As suggested by Manzano et al. (2008), economies with a significant oil industry have specific features that strongly influence how the institutional framework and the political economy of the sector evolve (for example, might need or have more fiscal discipline). Amundsen (2014) also warns us that rents from oil activity can be channeled into the productive economy or captured through institutional factors for reasons not necessarily driven by national goals. Furthermore, Ross (2001) and Martinez (2022) express how oil-rich regimes might manipulate financial information to reduce the threat of political turnover, creating environments of low or no taxation as well as extensive patronage networks. In this sense, do light emissions have the same predictive power on economic activity in oil regions? Can we assume similar measurement errors within oil-dependent countries? This chapter is a recent effort to answer these questions.

I follow a technical procedure to extract, inter-calibrate, harmonize, and correct satellite data of nighttime lights from two sources to cover the period 1992–2019. I construct a panel of lights and subnational indicators of economic activity for eight oil-dependent countries that group 40 percent of the total oil production: Brazil, Canada, Colombia, China, India, Mexico, Russia, and the United States. Considering different specifications, I verify that lights have predictive power on economic activity, with statistically different elasticities between oil regions and non-oil regions. I also confirm that oil regions have better-measured data, while nighttime lights improve the ability to estimate true economic growth. The elasticity results, together with other estimated parameters, contribute to any analysis on subnational GDP in countries where we could distinguish between oil and non-oil regions.

The application of satellite-recorded nighttime lights for tracking our behavior is not new. Croft (1978) used satellite pictures to associate light emissions with human activities, identifying as sources of emissions: the urbanization process, agricultural and pastoral fires, and gas flares at oil fields.[1] During the eighties, Welch (1980), Foster (1983), and Sullivan (1989) highlighted the relevance of using these images to analyze energy consumption patterns. Nevertheless, those studies faced a technical limitation relying on film strips. Starting the nineties, a digital archive of satellite imageries of nighttime lights was available, showing spatially stable light sources and taking into

[1]Gas flaring, as a byproduct of oil production, is a controlled burning of natural gas in the oil and gas industry.

account cloud covers. Based on these inputs, Elvidge et al. (1997), Sutton and Costanza (2002), Doll et al. (2006), Ghosh et al. (2010), Nordhaus and Chen (2012), and Henderson et al. (2012), verified the relationship between economic activity and lights, the last two studies stating a statistical framework to measure economic growth.[2]

The nighttime light approach to analyze economic activity is gaining ground and reliability due to technological improvements, and its advantage to examine local dynamics is more evident over time. Dai et al. (2017) recently estimate the spatial distribution of GDP at different regional levels contrasting light data from two imaging-resolution generations of sensor instruments. Bruederle and Hodler (2018) also use nighttime lights to estimate human development at a local level, Ferreira (2018) tracks regional socio-economic outcomes for Namibia, and Wang et al. (2019) calculate subnational GDP in Uganda.

The use of satellite data to estimate regional economic activity has advantages. It may act as an independent alternative to reduce measurement errors. For example, this chapter confirms how light data could complement observed growth data. It also allows to aggregate estimates to any geographic level rather than only to specific administrative divisions. A better understanding of this local dynamic would allow policy makers to have an alternative approach accounting for geographic locations and oil activity while setting national and subnational policies. Furthermore, considering these regional differences enable us to carry out further studies about oil contribution and to obtain disaggregated data in oil-dependent countries with no data at all.

The chapter is organized as follows. In Section 2.2, I describe the procedure to select the sample, light data's features, and the inter-annual calibration, harmonization, and top-coding correction carried forward. Section 2.3 examines stylized facts within and across the selected oil-dependent countries, giving a first approximation of the relationship between economic activity and nighttime lights by oil and non-oil regions. Section 2.4 describes the framework and presents the specifications and main results, while Section 2.5 concludes.

[2]The remote sensing approach to examine human economic activity does not circumscribe to the use of nighttime lights. For example, Burchfield et al. (2006) used land cover data to track the evolution of land use identifying the causes of urban sprawl in the United States, while Keola et al. (2015) used land cover data but to estimate the contribution of the agricultural sector to some developing economies.

2.2 Data

2.2.1 Sample Selection

Firstly, I focused on countries that produced oil at least once between 1992 and 2019 to categorize their oil and non-oil regions. Secondly, I ranked the countries' world share of cumulative oil produced throughout the period to obtain a rough selection of top producers. Finally, I identify which of those countries have official regional data of real GDP in local currency (LCU) or its equivalent (for instance, chain volume measures), having at least one growth rate calculation over the period. Table 2.1 shows the regional indicators of economic activity of the eight selected countries: Brazil, Canada, China, Colombia, India, Mexico, Russia, and the United States.[3] These countries group around 40 percent of the total oil produced worldwide.

Table 2.1: Regional Indicators of Economic Activity

Country	Sample	Indicator of regional economic activity	Source
Brazil	1998-2018	States gross value added (millions LCU, constant prices, preceding year)	Brazilian Institute of Geography and Statistics
Canada	1997-2019	Provincial gross domestic product (millions LCU, chained base year=2012)	Statistics Canada
China	1992-2019	Index of growth regional product (constant prices, preceding year=100)	National Bureau of Statistics of China
Colombia	1992-2019	Departmental gross domestic product (billions LCU, volume chained base year=2015)	National Department of Statistics
India	1992-2019	Gross state domestic product (crore LCU, constant prices base year=2011-2012)	Open Data initiative of Government of India
Mexico	1992-2019	Regional gross domestic product (millions LCU, constant prices base year=2013)	National Institute of Statistics and Geography
Russia	1998-2019	Volume index of gross regional product (constant prices, preceding year)	Federal State Statistic Service
United States	1997-2019	Real total gross domestic product (millions LCU, chained base year=2012, not seasonally adjusted)	U.S. Bureau of Economic Analysis

2.2.2 Nighttime Lights

Satellite data from two sources are extracted, inter-calibrated, and harmonized from 1992 to 2019. On the one hand, yearly satellite imagery of nighttime lights is available from the U.S. Air Force Defense Meteorological Satellite Program (DMSP) Operational Linescan System -version 4- from 1992 to 2013.[4] DMSP products are separated by satellite-year

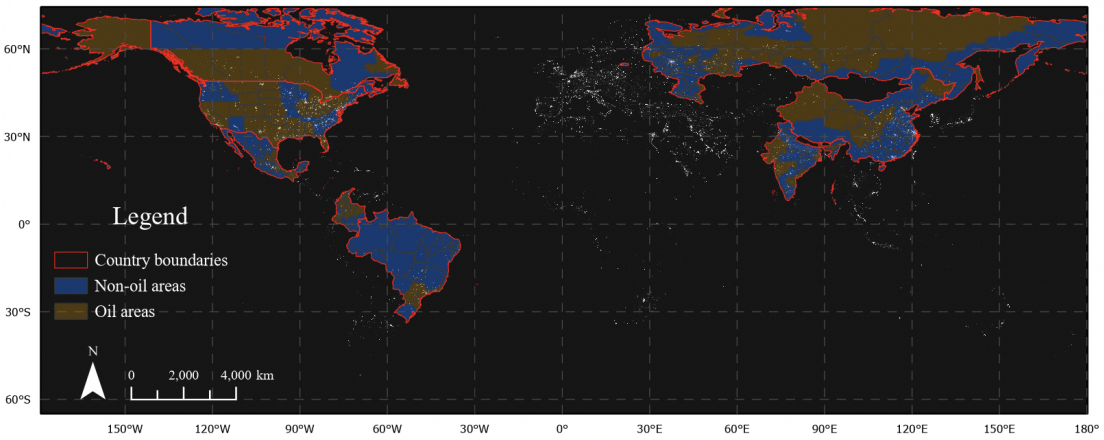
[3] This chapter refers to any subnational territory as "region" even though I am aware that countries have different administrative divisions: Brazil-States, Canada-Provinces, China-Provinces, Colombia-Departments, India-States/Union territories, Mexico-States, Russia-Federal subjects, and the United States-States. The detailed information is in Appendix A, Section A.1.

[4] There is DMSP imagery from satellite F10 and F12 for the year 1994, F12 and F14 for 1997–1999, F14 and F15 for 2000–2003, and F15 and F16 for 2004–2007. I use simple averages across satellites

in two sets of zipped files "average visible lights, stable lights, and cloud-free coverages" and "average lights adjusted by frequency of light detection." On the other hand, the Earth Observation Group of the Payne Institute for Public Policy (Colorado School of Mines) produces a new -version 2- consistently processed time series of annual global nighttime light imagery from the Visible Infrared Imaging Radiometer Suite (VIIRS) on-board the Suomi National Polar-Orbiting Partnership satellite platform available for the period 2012–2019. Looking for pre-filtered data, I use the DMSP annual composites of stable lights and cloud-free coverages and the VIIRS annual masked average radiance products (Elvidge et al. 2021).[5]

Each satellite imagery contains information about the intensity of lights on a grid. For example, Figure 2.1 shows the 2019 spatial distribution of light emissions from the VIIRS annual composite, highlighting the selected countries, oil-producing regions (brown), and non-oil producing regions (blue). In this case, more light intensity relates to whiter pixels, while unlit areas are black.

Figure 2.1: VIIRS Nighttime Lights: Selected Sample, 2019



Source: Earth Observation Group, Payne Institute - Colorado School of Mines.

There is heterogeneity among countries. China, Russia, and the United States show a higher light density throughout their coastal areas. Canada has most of its lights concentrated within oil-producing regions, which is associated with around 44 percent of

within pixel-years (Henderson et al. 2012).

[5]The files are at the online repository of the Earth Observation Group (2021, October): https://payneinstitute.mines.edu/eog/nighttime-lights/. For details, see the readme file associate with the respective sources.

its total land area. Colombia and Brazil have a similar pattern to Canada, but with a significant difference. The oil region covers nearly 58 percent of Colombia's land area, while it represents only 6 percent for Brazil. India and Mexico reveal more uniformity in light sparsity, but Mexico has most of its lights surround its Gulf and the capital.

Table 2.2 compares the main features of light data from DMSP and VIIRS. In the case of DMSP, the intensity of light is measured as a six-bit digital number (DN), recorded for every 30 arc-second output pixels (i.e., around 0.86 square kilometers at the equator). Unlit areas -as well as clean areas from background noise-take values of zero, while the rest of the data lie in the range between 1 and 63 (in which 63 expresses the brightest light). DMSP sensors lack lineality and have no on-board calibration, which means that the DN values assigned to pixels are not comparable from one year to another. This situation hampers monitoring changes in lighting at national and subnational levels (Hsu et al. 2015; Jeswani et al. 2019; Levin et al. 2020; Ma et al. 2020). Regarding VIIRS, its spatial resolution provides more detail than DMSP. In addition, VIIRS has on-board calibration and a wider radiance range leading to more robust low light detection, thus not suffering from frequent saturation as in DMSP imagery.

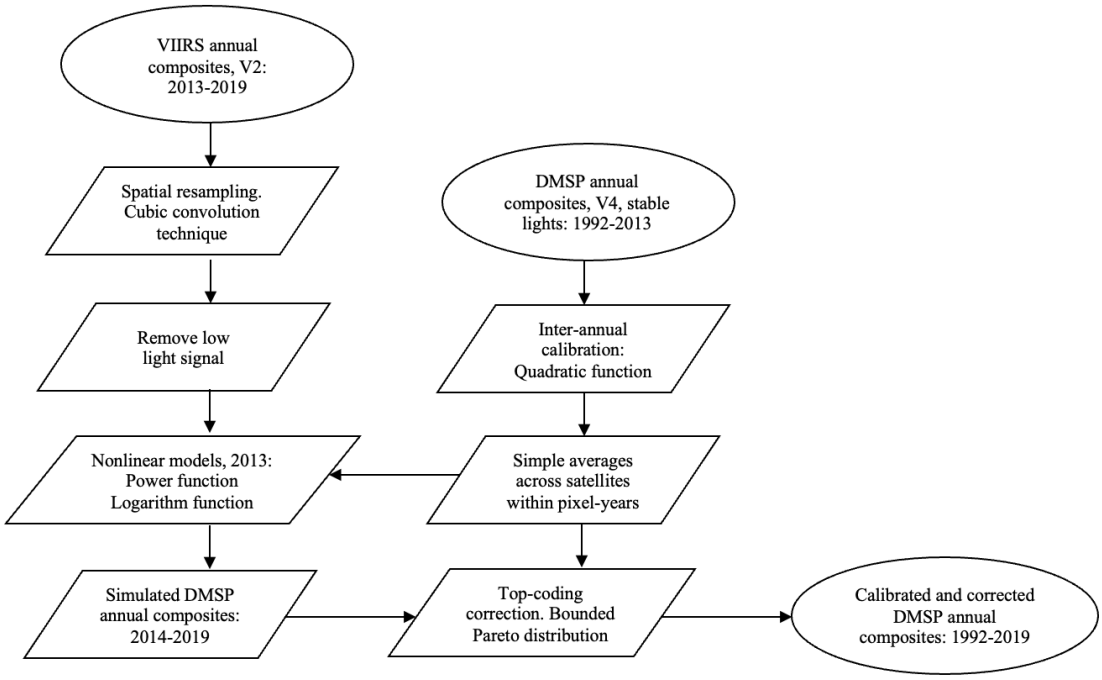
Table 2.2: Comparison of Features Between DMSP and VIIRS

	DMSP	VIIRS
Spatial resolution	30 arc-second	15 arc-second
Radiometric resolution	6-bit	14-bit
Low light imaging bandpass	0.5-0.9 μm	0.5-0.9 μm
Nighttime overpass	~19:30	~1:30
On-board calibration	No	Yes
Units of pixel values	0-63 scale	Radiance (nanoWatts/cm^{2}sr)
Light range detected (Wcm^{-2}sr$^{-1}\mu^{-1}$)	1.54x10^{-9}-3.17x10^{-7}	3x10^{-9}-0.02 (specified, actual detected noise floor 5x10^{-11})
Saturation	Common in urban cores	No
Availability (free)	1992-2013 annual composites	2012-present monthly composites 2012-present annual composites

The differences between DMSP and VIIRS are notable (Elvidge et al. 2013; Sánchez de Miguel et al. 2014; Li et al. 2017; Jeswani et al. 2019; Wu and Wang 2019; Levin et al. 2020, Sánchez de Miguel et al. 2021). Li et al. (2017) mention these differences are motivated by sensor variations in spatial resolution, spectral response, point of spread function, overpass time at night, and wider radiance range of the VIIRS. Therefore, analyzing the timeframe 1992–2019 requires a pre-processing of light input data to

extend DMSP annual composites throughout 2014–2019.[6] Figure 2.2 summarizes the main steps.

Figure 2.2: Pre-Processing of Nighttime Light Products



2.2.3 Pre-Processing: Inter-Annual Calibration

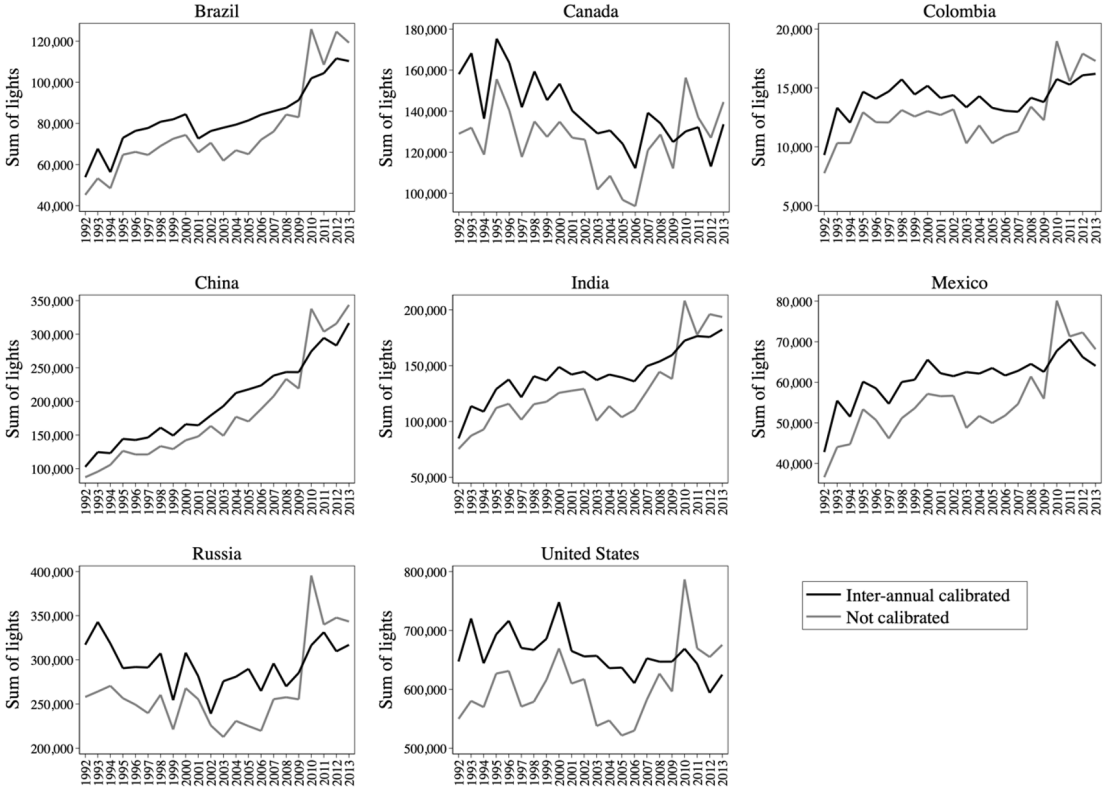
The lack of on-board calibration of the DMSP sensor requires carrying out an annual inter-calibration procedure across all the available stable light products (Elvidge et al. 2009; Hsu et al. 2015; Jeswani et al. 2019; Ma et al. 2020).

In particular, this chapter uses the method proposed by Elvidge et al. (2009) to improve the DMSP data from 1992 to 2013. They suggested a flexible framework for calibrating DMSP lights. They identified Sicily, Italy, as a place where the sum of lights had not significantly changed across time and chose the satellite-year of its maximum light intensity as a benchmark to derive coefficients from a second-order polynomial regression. The idea was to reduce the volatility of total light value between different years and satellites.

[6] For details, see Subsection 2.2.3, Subsection 2.2.4, and Appendix A, Section A.2.

The inter-calibration consists of using the satellite-specific coefficients from the quadratic function to obtain fitted DN values for all the DMSP sample. After the inter-annual calibration, a within-pixel simple average was calculated in each year with DMSP data from two satellites. Overall, this procedure enhances the continuity of DMSP data and reduces its fluctuations.[7] Figure 2.3 shows the results.

Figure 2.3: Inter-Annual Calibrated versus Not Calibrated DMSP Lights by Country



Source: Own calculations.

2.2.4 Pre-Processing: Harmonization

The spatial resolution between the sensors is different. VIIRS annual composites was first aggregated to the spatial resolution of DMSP using cubic convolution as resampling operator to then extract DMSP and VIIRS data averaging values from grid cells of size 10×10 km. Following Li et al. (2017) and Wu and Wang (2019), 0.3 nanoWatts/cm2sr is

[7] For details, see Elvidge et al. (2009), Hsu et al. (2015), and Ma et al. (2020).

fairly assumed as the threshold for low signal values out of the detecting range of DMSP sensors, thus I subtract it from VIIRS annual composites (if negative, the value is set to zero).

From 2013 to now, VIIRS annual pre-processed data accounts for all the months, but the 2012 composite used monthly data from April to December. Zhao et al. (2019) suggest using the relationship of 2013 between VIIRS and DMSP to reduce the impact of seasonal fluctuations caused by the original VIIRS data while simulating 2014–2019 DMSP data. Therefore, I use the common year, 2013, to estimate a logarithm and a power function specification of DMSP on VIIRS, using the root mean square error (RMSE) to identify the best fit.

The logarithm specification is suggested to address the curvature when the DN fitted values approached saturation in the inter-annual calibrated composite, while the power function account for areas with overall values far from the saturation level.

The estimated parameters from the non-linear regressions are used to construct DMSP annual composites from 2014 to 2019. The values above a threshold of 63 are restricted to be equal to 63 DN as a way to obtain similar units.[8] Finally, the analysis is based on 23,492,556 pixels or 7,371 observations at the regional level.

2.2.5 Pre-Processing: Top-Coding Correction

The data suffers from right-censoring originated by sensor saturation. Most of this literature leaves aside the top-coding issue arguing that the fraction of pixels equal to 63 DN is near zero or zero in low and middle-income countries, and that even in very densely populated cities in high-income countries, the share is still small.[9] Nevertheless, this is not necessarily the case.

As indicated by Bluhm and Krause (2018), light data have already been averaged or smoothed during raw construction at least twice. Therefore, pixels with lower values should be brighter than recorded.[10] Consequently, top-coding is not only present at 63 DN. Hsu et al. (2015) also warn us about the presence of saturation even in the mid-30s. This situation could be especially true in this analysis. On the one hand, the data are

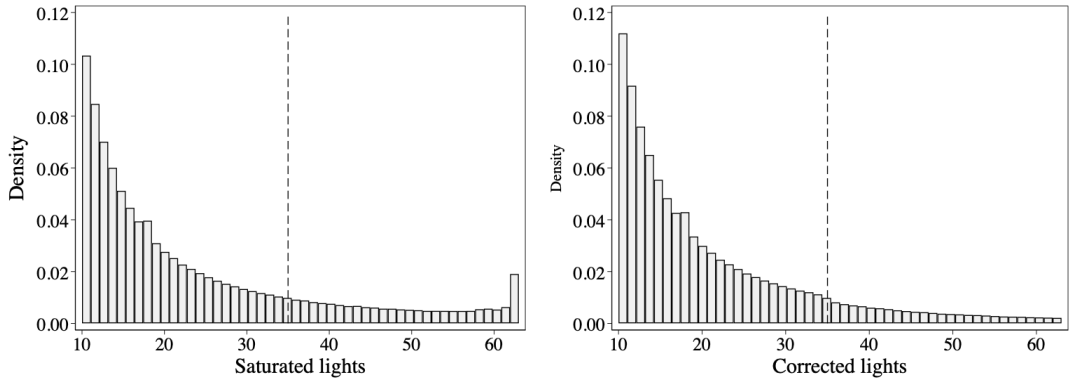
[8] For details, see Appendix A, Section A.2.

[9] For instance, see Henderson et al. (2012) They mentioned the case of the Netherlands in which they found that only 1.58 percent of pixels were top-coded.

[10] See Abrahams et al. (2018) for a detailed description of how DMSP satellites process the data.

processed on a pixel area of 10x10 km, obtaining even more smoothed data than those raw values. On the other hand, the data are analyzed at the subnational level. A dense city of a high-income country might represent a low fraction of top-coding relative to its country, but most of its pixels may suffer saturation. Therefore, top-coding might be an issue. Figure 2.4 shows the distribution of light intensity before and after the correction.

Figure 2.4: Histograms of Saturated versus Corrected Lights (All Years)



Source: Own calculations. Note: The density is calculated throughout all the values of light intensity but, for practical reasons, I only show from 10 DN to 63 DN. On the left panel, the distribution before the correction. On the right panel, the distribution after the correction (the highest value is 1,999). The vertical dash line denotes the top-coding threshold.

The correction approach follows Bluhm and Krause (2018). They propose a truncated (bounded) Pareto distribution as a reasonable description of top lights suggesting a shape parameter of 3/2. In this chapter, I use Bluhm and Krause (2018)'s time-invariant shape parameter across all countries, with a saturation threshold set to 35 DN and an upper bound to 2000.[11] Finally, to obtain non-censored observations, I apply the inverse-transform method over the bounded distribution and use a country-year random draw from a standard uniform distribution.[12]

[11] Bluhm and Krause (2018) identify 2,000 as a value close to the average maximum observed in the brightest cities in the world.

[12] For details, see Appendix A, Section A.3.

2.3 Nighttime Lights and Economic Activity: Heterogeneity among Countries-Regions

The analysis uses calibrated, harmonized, and corrected nighttime light data. Table 2.3 summarizes some statistics associated with the distribution of nighttime lights and population.[13]

Table 2.3: Summary of Nighttime Lights and Population by Region (Average)

	Brazil		Canada		China		Colombia		India		Mexico		Russia		United States	
	Non-oil	Oil	Non-oil	Oil	Non-oil	Oil	Non-oil	Oil	Non-oil	Oil	Non-oil	Oil	Non-oil	Oil	Non-oil	Oil
Area unlit	81.0%	27.0%	96.2%	78.6%	61.5%	75.8%	91.4%	65.2%	32.6%	14.8%	55.6%	28.6%	78.8%	87.8%	30.9%	60.2%
$0 < DN < 1$	6.8%	16.5%	0.7%	4.9%	10.2%	7.2%	2.4%	10.6%	13.6%	11.3%	13.4%	17.0%	6.1%	2.9%	8.4%	7.3%
$1 \leq DN < 3$	7.0%	23.5%	0.5%	3.6%	9.5%	6.4%	1.8%	12.6%	14.3%	17.2%	15.0%	19.0%	5.0%	2.1%	10.3%	8.9%
$3 \leq DN < 6$	2.5%	11.4%	1.2%	6.1%	8.4%	5.1%	1.8%	5.2%	16.5%	31.8%	6.8%	18.0%	4.9%	3.7%	17.1%	8.0%
$6 \leq DN < 11$	1.5%	9.0%	0.6%	4.4%	4.9%	2.7%	1.2%	3.5%	14.4%	17.1%	4.7%	9.7%	3.3%	2.4%	15.1%	6.7%
$11 \leq DN < 35$	1.1%	10.1%	0.8%	2.1%	4.3%	2.4%	1.3%	2.5%	8.1%	7.1%	3.7%	6.6%	1.6%	1.0%	13.7%	6.7%
$35 \leq DN < 63$	0.1%	1.5%	0.1%	0.3%	0.7%	0.3%	0.1%	0.3%	0.4%	0.4%	0.5%	0.9%	0.2%	0.1%	2.7%	1.3%
$63 \leq DN$	0.1%	1.1%	0.1%	0.2%	0.5%	0.1%	0.1%	0.2%	0.3%	0.2%	0.4%	0.3%	0.1%	0.1%	1.9%	0.9%
DN (average per square kilometers)	0.71	6.39	0.38	1.45	2.91	1.35	0.57	1.67	4.18	4.90	2.51	3.91	1.13	0.75	8.74	4.51
Population density (average per square kilometers)*	12	99	1	3	134	56	15	46	302	271	42	62	6	2	28	16
Gini index for lights (all country)	0.7193		0.6380		0.6953		0.6734		0.5294		0.6917		0.6252		0.6824	
Gini index for lights	0.6927	0.7234	0.6264	0.6351	0.7048	0.6687	0.6552	0.6727	0.5619	0.4776	0.7136	0.6160	0.6466	0.5965	0.6461	0.7015
Gini index for population (all country)*	0.8961		0.9399		0.8160		0.9012		0.6235		0.8971		0.9197		0.8979	
Gini index for population*	0.8773	0.8769	0.9230	0.9448	0.7834	0.8508	0.9403	0.8763	0.6212	0.6215	0.9165	0.7908	0.9040	0.9313	0.8617	0.9133

Note: DN should be an integer, but I extracted the data on a pixel area of ten square kilometers. Pixel values of zero are excluded to calculate the Gini index. *Average from 2000 to 2019.

There are differences among countries and their regions. The non-oil areas of India and the United States are unlit at about 32.6 percent and 30.9 percent, respectively. That is expected from India, with the highest population density among all the countries (11 times more than the density in the United States). With more population density than the United States, Mexico and China have more non-oil areas with no lights (55.6 percent and 61.5 percent, respectively). The non-oil regions of Colombia and Brazil have a similar population density than the United States' oil region, but the proportion of areas unlit is higher in those countries. This situation might express the differences in income levels, especially accounting for how sparse the population is.

Regarding the oil-producing regions, Brazil, India, and Mexico have less than 30 percent of their territory unlit, from which India has more population density sparsed

[13]The repository of the Oak Ridge National Laboratory (ORNL) provides LandScan global population distribution data at approximately 1 km resolution (30 arc-second), in which the values of the cells are counts from an ambient population (average over 24 hours). With this dataset, it is possible to relate the population with nighttime lights at a pixel level.

through pixels. Canada and China have similar values for area unlit (around 75 percent or more), but the population density of Canada is significantly lower and close to Russia (which has over 85 percent of unlit pixels). Independently of the country or region in the sample, the fraction of corrected pixels (equal to or above 35 DN) is low. Nevertheless, Brazil has almost 1.1 percent of corrected lights in oil regions, and the United States has nearly 2 percent in its non-oil region.

In the aggregate, following the Gini coefficients, the more significant gap is for India, in which population and lights are more disperse across pixels. The gap in the populations spatial distribution relative to lights tends to be greater within oil regions. The spatial distribution of light also appears to vary through the sample. Therefore, it is necessary to account for heterogeneity.

It is expected that rural regions relate to less illuminated areas, which could explain the low intensity of lights in areas with vast land. This circumstance is a limitation of the methodology. Keola et al. (2015) warn us that agricultural activity is better estimated through land cover observations rather than lights independently of the degree of development of a country. They conclude that agricultural activity emits marginal lights -if any. Hence, the assumption that light emissions explain economic activity bases on the idea that urban areas are the "enclave sector" of the economy, leaving agriculture somewhat implicitly defined in the measurement error. Further studies should consider this limitation, especially those analyses focused on developing countries (where agriculture plays an important role).

The oil and gas industry tends to set in rural areas. This industry is mainly capital rather than labor-intensive, and most of the land surrounding gas flares may have low population density. If this is the case, the relationship between population data, lights, and economic activity might not be evident. Therefore, why could we have a significant amount of light emissions in oil regions?

One explanation could be related to the identification of oil regions. I identify top producer countries, not top producer regions within countries. There is a trade-off between an extended sample of oil producer countries (with different degrees of development) and oil production data. The latter is not a trivial task. Either not all countries centralize their oil accounts, or high oil-dependent countries use those accounts as a sensible policy topic. Nevertheless, this chapter's main focus is to verify if there are

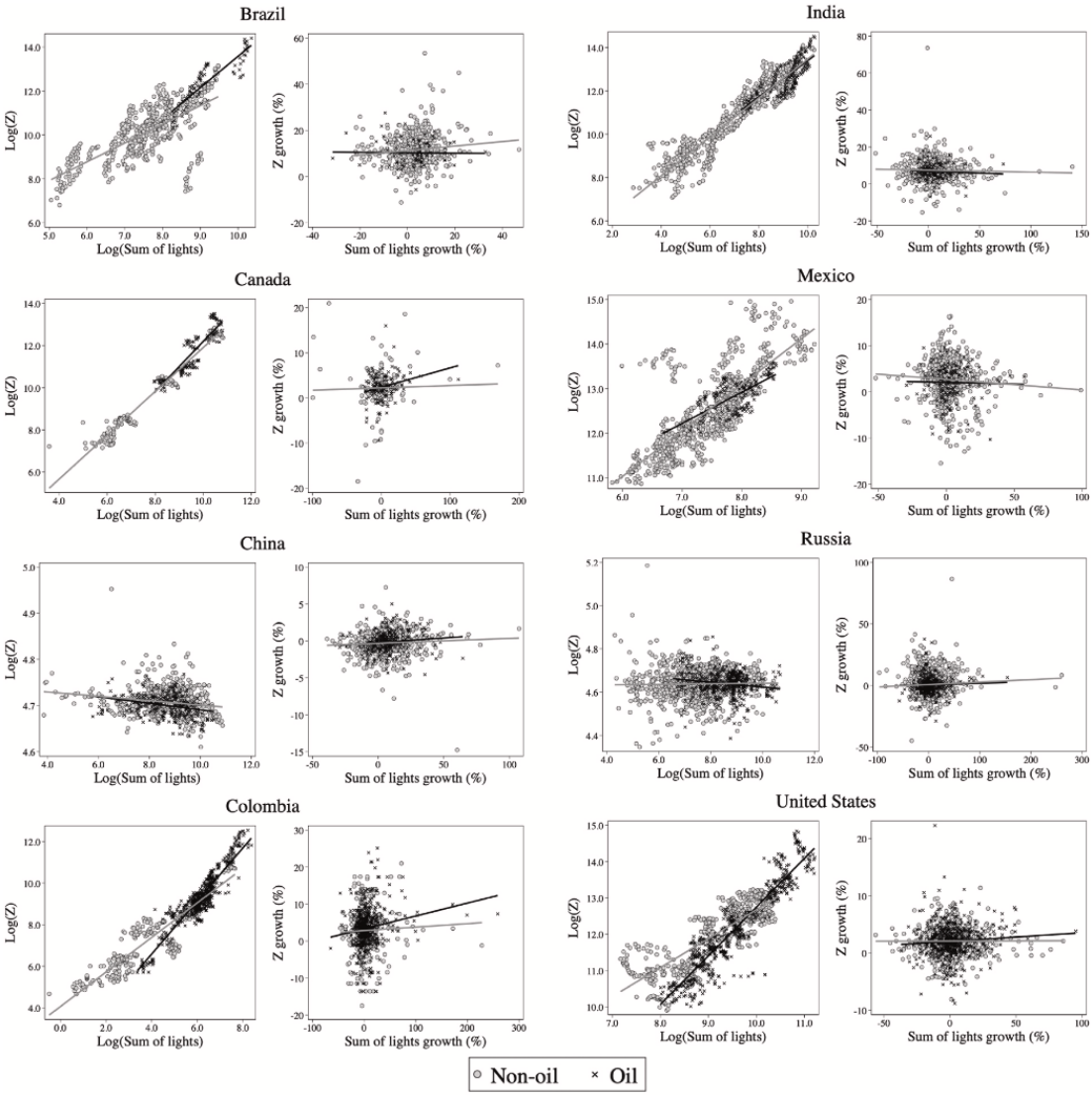
differences in economic activity and nighttime lights by region, for which oil production series are not needed. Furthermore, following national and international sources listing a particular area as an oil producer strengthens the identification process.[14] A second explanation is the presence of satellite-recorded gas flaring.[15] DMSP products do not discriminate among sources of lights. Most studies treat gas flares as background noise, but gas flaring correlates positively with oil production (Debbich 2019). Therefore, it might represent a valuable source of information, especially for oil-producing countries. For this reason, I use lights data with gas flaring. A final explanation should be clear: the intensity of light reveals different sources of economic activities, not only from oil production.

Figure 2.5 is a first approximation to identify any pattern between lights and economic activity by type of region. For most countries, a positive linear relationship arises from region-level data. Only China and Russia do not seem to show a clear pattern. The rest of the countries show more concentration of oil-producing regions to the right of their plots relative to non-oil areas; therefore, oil regions may be associated with more light intensity within their areas. Moreover, some countries denote that concentration in the upper-right of their plots. As expected, it is more challenging to find *a priori* a clear pattern in growth rates, but, even in this case, there is a slightly positive relationship between lights and economic growth for some countries. Henceforth, the analysis would be carried out on the data at levels.

[14]The ideal is to have regional oil production by year to identify any binary shift-producer or not, particularly for areas with low oil activity. Still, as Manzano et al. (2008) mentioned, the oil industry is related to high-sunk costs, which could lead to rigidities once a region starts to develop oil areas.

[15]Offshore gas flaring cannot be associated with a particular area within a country but with the country as a whole. Thus, I focus the attention on lights on the land surface.

Figure 2.5: Economic Activity and Sum of Lights by Country-Region (All Years)



Source: Own calculations. Note: Z represents the economic activity indicators from Table 2.1. To reduce outlier's dominance due to lights output, I exclude: on the left panels, two observations for Canada (Northwest Territories, 2003; Nunavut, 2003); on the right panels, three observations for Canada (Northwest Territories, 2004; Nunavut, 2002 and 2004), and five observations for Russia (Chukot, 2004; City of St. Petersburg, 2003; Murmansk, 2000; Samara, 2014; Tuva, 2014).

2.4 Estimation and Results

2.4.1 Specifications

Different specifications have been used to estimate economic activity based on nighttime lights. For example, using aggregated data, Henderson et al. (2012) state a classical measurement error framework in economic growth to specify a production relationship between light growth and income growth.

In their context, let y be the growth in 'true' real GDP (with no measurement errors), z the growth of real GDP as measured in national income accounts, and x the growth of observed light. They assume that the GDP growth of each country j, as recorded in national income accounts, may suffer from classical measurement error (Equation 2.1), with a variance of true income growth represented by σ_y^2 and a variance of ϵ_z by σ_z^2.

$$z_j = y_j + \epsilon_{z,j} \tag{2.1}$$

They suggest a relationship between the growth of lights and the growth of true income as given by Equation 2.2, where β is a "structural" parameter indicating the elasticity of lights with respect to income, while the variance of ϵ_x could be represented by σ_x^2. The error term in this equation could be understood as a noise in how measured light growth reflects GDP growth (for example, representing a measurement error in lights, that is, the difference between what true light reflects into space and what the satellite captures or due to variation among countries in the relationship between GDP growth and growth of light emissions).

$$x_j = \beta y_j + \epsilon_{x,j} \tag{2.2}$$

In this sense, the measurement error in GDP is not necessarily related to the error in the equation determining observable light; therefore, they assume that $cov(\epsilon_x, \epsilon_z) = 0$.

Only for predictive purposes, Henderson et al. (2012) also propose a regression of growth of income on growth of lights, such in Equation 2.3.

$$z_j = \hat{\Psi} x_j + e_j \tag{2.3}$$

In their case, the coefficient $\hat{\Psi}$ is obtained through a fixed-effect model in an Ordinary Least Squares (OLS) regression, where $\hat{\Psi} = \frac{cov(x,z)}{var(x)}$. Assuming that the degree of measurement error in the economic growth does not affect the estimated value of $\hat{\Psi}$ allows them to obtain a best fit relationship to be used in producing proxies for income growth, $\hat{z_j} = \hat{\Psi} x_j$.

For this chapter, their framework could be valuable. I use subnational rather than aggregated data, as well as calibrated, harmonized and top-coding corrected lights. Also, it would be helpful to verify if there are local differences between oil and non-oil regions in how lights could predict economic activity and how the official data are reported. They also suggest an extension of the approach considering the presence of more than one group (shortly explained below).[16]

They argue that if we do not observe local economic growth data, a predictive parameter $\hat{\Psi}$ and Equation 2.3 could be used to obtain proxies. However, using observed data improves the precision of estimated growth in reflecting the estimation of true growth.

Let y now be the true behavior of the economic activity (true growth), k an indicator of regions as oil producers or not, t years, and j regions as defined by the indicator. It is possible to improve the precision of estimated growth by combining a separate estimate of economic growth using light data with an observed measure of economic growth. Equation 2.4 represents the composite estimate of growth, $\hat{y}$, by specifying weights that minimize the variance of measurement error in this estimate relative to the true growth.

$$\hat{y}_{tj,k} = \lambda_k z_{tj,k} + (1 - \lambda_k)\hat{z}_{tj,k} \tag{2.4}$$

Based on their assumptions about error structure, the next expression shows the variance of this composite estimate (subscripts are suppressed for now).

$$var(\hat{y} - y) = var(\lambda(z - y) + (1 - \lambda)(\hat{z} - y)) = \lambda^2\sigma_z^2 + (1 - \lambda)^2 var(\hat{z} - y) \tag{2.5}$$

The last term in this expression can be expanded as follows:

[16] For details, see Henderson et al. (2012).

$$var(\hat{z} - y) = var(\hat{\Psi}x - y) = var(\hat{\Psi}\beta y + \hat{\Psi}\epsilon_x - y) = (\hat{\Psi}\beta - 1)^2\sigma_y^2 + \hat{\Psi}^2\sigma_x^2$$

Noting that the relationship between $\hat{\Psi}$ and the structural parameter β can be represented as $plim(\hat{\Psi}) = \frac{1}{\beta}\left(\frac{\beta\sigma_y^2}{\beta^2\sigma_y^2+\sigma_x^2}\right)$, the expression could be rewritten as:

$$var(\hat{z} - y) = \frac{\sigma_y^2\sigma_x^2}{\beta^2\sigma_y^2 + \sigma_x^2}$$

Substituting this into Equation 2.5:

$$var(\hat{y} - y) = \lambda^2\sigma_z^2 + (1 - \lambda)^2\frac{\sigma_y^2\sigma_x^2}{\beta^2\sigma_y^2 + \sigma_x^2} \tag{2.6}$$

Then, it is possible to use Equation 2.6 to solve for the optimal weight λ^* which minimizes this variance by type of region, obtaining Equation 2.7.

$$\lambda_k^* = \frac{\sigma_x^2\sigma_y^2}{\sigma_x^2\sigma_y^2 + \sigma_{z,k}^2(\beta^2\sigma_y^2 + \sigma_x^2)} \tag{2.7}$$

There are only unknown parameters on the right side, so they derive auxiliary equations to solve using the three sample moments provided by the observed data (Equation 2.8-Equation 2.10) and assuming a ratio of signal to total variance in measured z for oil and non-oil regions, defined in Equation 2.11.

$$var(z) = \sigma_y^2 + \sigma_z^2 \tag{2.8}$$

$$var(x) = \beta^2\sigma_y^2 + \sigma_x^2 \tag{2.9}$$

$$cov(y, x) = cov(x, z) = \beta\sigma_y^2 \tag{2.10}$$

$$\phi_k = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_{z,k}^2} \tag{2.11}$$

I assume that $\sigma_{z,k}$ depends on whether the regions are oil producers or not, although

imposing the same light-economic structure for them. Specifying a value of signal in Equation 2.11 for the set of oil-producing regions allows to use Equation 2.8-Equation 2.11, to then solve Equation 2.7.

Other studies estimate the dynamic of economic activity using data at level. For example, Dai et al. (2017) use provincial and city-level GDP data of China under three specifications: a linear regression, a power function, and a polynomial model. They collapse the sum of the lights from the grid, to then use it as a regressor explaining the real GDP. In this case, they conclude that the results based on different models do not have significant differences; therefore, a linear model might be used.

Following these ideas, for predictive purposes, I propose at least four logarithmic specifications over an unbalanced panel at the regional level j with years t.[17]

$$Log(Z)_{tj} = \alpha Log(Sum\ of\ lights)_{tj} + c_j + d_t + v_{tj} \tag{2.12}$$

$$Log(Z)_{tj} = \theta Log(Intensity\ of\ lights)_{tj} + c_j + d_t + u_{tj} \tag{2.13}$$

$$Log(Z)_{tj} = \beta_1 Log(Intensity\ of\ lights)_{tj} + \beta_2 Log(Intensity\ of\ lights)^2_{tj} + c_j + d_t + w_{tj} \tag{2.14}$$

$$Log(Z)_{tj} = \gamma Log\left(\frac{Sum\ of\ lights}{Population}\right)_{tj} + c_j + d_t + \epsilon_{tj} \tag{2.15}$$

The indicator for economic activity is Z. In their most basic structure, Z is regressed on the sum of lights as in Equation 2.12, on the regional average of lights per square kilometers or intensity of lights as in Equation 2.13, on the intensity of lights including its quadratic term as in Equation 2.14, and on the sum of lights adjusted by population as in Equation 2.15. Each specification includes country and year fixed effects (c and d, respectively), as well as clustered standard errors by regions.

[17] This allows to reach estimated parameters expressing elasticities, attenuating problems of heteroskedasticity and asymmetric distributions of errors.

2.4.2 Results

Table 2.4 shows the results for the basic setup.[18] In all cases, the nighttime lights have a significant statistical effect predicting economic activity. Furthermore, the quadratic specification cannot be immediately discarded, indicating possible nonlinearities between lights and economic activity as the light intensity scales-up. This result is aside from the correction of the light data.

Table 2.4: Estimations: All Regions

	(1) $Log(Z)$	(2) $Log(Z)$	(3) $Log(Z)$	(4) $Log(Z)$
$Log(Sum\ of\ lights)$	0.6722***			
	(0.0445)			
$Log(Intensity\ of\ lights)$		0.2882***	0.2977***	
		(0.0466)	(0.0419)	
$Log(Intensity\ of\ lights)^2$			-0.0389***	
			(0.0101)	
$Log(Sum\ of\ lights/Population)$				-0.1592*
				(0.0931)
R^2 (within)	0.5809	0.2152	0.2608	0.0286
Observations	7,369	7,369	7,369	5,986
Groups	8	8	8	8

Note: Significant at *10, ***1 percent. Clustered standard errors adjusted for 305 regional groups in (). The inclusion of population restricts the sample to years after 2000, inclusive.

On the other hand, Column (4) explores the effect of spatial dissimilarities between lights and population.[19] In the aggregate, lights adjusted by population density are only statistically relevant explaining economic activity outputs at 10 percent of significance level, which could signal overlapping effects between them.

Table 2.5 reveals results from extended versions of each specification (i.e., including an indicator variable for oil-producing regions and interaction terms). Column (1) confirms similarities between oil and non-oil regions on how their sum of lights explains the economic activity. However, that result differs from using the intensity of light, Column (2) and Column (3), which accounts for each administrative region's total size. In those

[18] A summary of some statistics for the different variables is presented in the Appendix A, Section A.4.

[19] A less restrictive way to do this is to use regional GDP per capita as a dependent variable. However, regional GDP was not available for all the countries.

cases, in this chapter I verify the presence of statistical heterogeneities. This finding suggests the need to start accounting for subnational particularities in oil-producing countries. In Column (4) lights adjusted by population also gains statistical significance in oil regions.

Table 2.5: Estimations: Statistical Differences by Region

	(1) $Log(Z)$	(2) $Log(Z)$	(3) $Log(Z)$	(4) $Log(Z)$
$Log(Sum\ of\ lights)$	0.6659***			
	(0.0463)			
$Log(Sum\ of\ lights)$ x Oil	0.0299			
	(0.0724)			
$Log(Intensity\ of\ lights)$		0.2586***	0.2669***	
		(0.0465)	(0.0431)	
$Log(Intensity\ of\ lights)$ x Oil		0.1576**	0.1522**	
		(0.0756)	(0.0800)	
$Log(Intensity\ of\ lights)^2$			-0.0333***	
			(0.0107)	
$Log(Intensity\ of\ lights)^2$ x Oil			0.0383	
			(0.0392)	
$Log(Sum\ of\ lights/Population)$				-0.1399
				(0.0938)
$Log(Sum\ of\ lights/Population)$ x Oil				-0.3029***
				(0.1062)
Oil	-0.2325	0.5034***	0.3002**	-1.3507*
	(0.6121)	(0.1271)	(0.1312)	(0.7100)
R^2 (within)	0.5813	0.2928	0.3212	0.1367
Observations	7,369	7,369	7,369	5,986
Groups	8	8	8	8

Note: Significant at *10, **5, ***1 percent. Clustered standard errors adjusted for 305 regional groups in (). The inclusion of population restricts the sample to years after 2000, inclusive.

In general, the elasticity for oil regions is higher than for non-oil regions. More predictive power in oil regions is also confirmed by increases in goodness of fit, which are revealed across Column (2)–Column (4) in Table 2.5. The use of the intensity of lights rather than the sum of lights confirms the difference between regions (the interaction is statistically significant). Considering the intensity of light, the elasticity for non-oil regions is about 0.2586 versus 0.4162 for oil regions. These elasticities are similar to those from the quadratic regression.

Table 2.6 summarizes statistical differences in variance between the regions. The

results indicate the output after regressing the squared residuals of each specification in Table 2.5 on a constant term and the oil-indicator variable.

Table 2.6: Squared-Residuals Regressions

	(1) $\hat{v}^2$	(2) $\hat{u}^2$	(3) $\hat{w}^2$	(4) $\hat{\epsilon}^2$
Oil	0.0113	-0.1992***	-0.1751***	-0.0904*
	(0.0216)	(0.0335)	(0.0336)	(0.0485)
Observations	7,369	7,369	7,369	5,986

Note: Significant at *10, ***1 percent. Clustered standard errors adjusted for 305 regional groups in (). The inclusion of population restricts the sample to years after 2000, inclusive. The residuals are those from Equation 2.12–Equation 2.15.

It is evident the negative sign on the squared-residuals regressions associated with light intensity variables. The oil indicator is statistically relevant in both regressions, reconfirming dissimilarities by region. The negative values mean that the volatility of the economic activity is greater for non-oil regions with respect to oil regions, that is, $\sigma^2_{z,k=0} > \sigma^2_{z,k=1}$ (with $k = 1$ if oil region). Therefore, allowing for measurement errors in economic growth implies oil regions with better-measured data than non-oil regions. Under this idea, I show Table 2.7.

Table 2.7: Results for Measured Economic Activity

Signal to total variance (ϕ_k)		Optimal weight (λ^*_k)	
Oil	Non-oil	Oil	Non-oil
1.0000	0.8689	1.0000	0.8541
0.9000	0.7820	0.8868	0.7574
0.8000	0.6951	0.7736	0.6608
0.7000	0.6082	0.6604	0.5641
0.6000	0.5213	0.5472	0.4674
0.5000	0.4345	0.4340	0.3707

Based on Henderson et al. (2012), I consider different scenarios of signal to total variance ratios for oil regions, to then use the equations previously mentioned to solve the parameters. To simplify and follow the presence of statistical differences, I assume $var(z_k)$ to be the variance of $z_{tj,k} - \hat{\phi} x_{tj,k}$, using the results from Table 2.5, Column (2).

Firstly, it is to highlight that both weights imply positives $1 - \lambda^*_{k=0}$ and $1 - \lambda^*_{k=1}$, which means that the use of nighttime lights improves the ability to measure true

economic growth in both types of regions. Secondly, it is noticeable how less variance on true growth is associated with a sharper decline in both weights. Therefore, lights would be contributing even more to estimate the true behavior of economic activity under less volatility. Thirdly, in an extreme scenario in which lights have no improvement measuring true growth in oil regions, it is still viable to improve the estimates in non-oil regions.

2.5 Conclusion

Local dynamic matters. The predictive power of nighttime lights on economic activity is verified at the regional level, with statistical differences between oil regions and non-oil regions. The relationship is always positive under different specifications, but it gains robust significance using light intensity rather than the sum of lights or adjusting by population. The quadratic component for light intensity cannot be entirely discarded even when using top-coding corrected lights. It denotes a negative and significant relationship with the economic activity, indicating possible nonlinearities when the light intensity scales-up, especially in non-oil regions. These results suggest the need to start accounting for subnational particularities in oil-producing countries.

It is also confirmed that light data improves the ability to measure optimal or true -with no measurement error- economic growth in both types of regions. Still, not in the same way. Oil regions seem to have better-measured regional data, and in an extreme scenario in which light data have no improvement measuring true growth in oil regions, it is still viable to improve the estimates using lights in the non-oil regions. Under less volatility on the economic activity, nighttime lights would be contributing even more to estimate its optimal behavior.

Chapter 3

Living in Darkness: Rural Poverty in Venezuela

3.1 Introduction

Today, most of the world's extreme poor live in rural areas. Rural poverty represents 80 percent of total extreme poverty, while about 45 percent of rural residents are at least moderately poor (Castañeda et al. 2018; Olinto et al. 2013). Most of them have no access to electricity, being electricity deprivation a category of "multidimensional poverty."[1] Despite significant efforts to increase global access to electricity,[2] there are still more than 750 million people suffering lack access to electricity, not equally distributed between rural and urban areas (Ritchie and Roser 2019). International Energy Agency et al. (2022) estimate that half of people without electricity live in fragile and conflict-affected settings and about 83 percent are in rural areas. The scenario is not different for Central and South America, where more than 25 million people lagged on electricity access. This data is significantly valuable on aggregated terms. However, deepening any analysis about poverty at the local level requires obtaining data overcoming availability issues from infrequent and expensive household surveys.

This chapter applies a remote sensing approach to address the unavailability of local rural poverty data for Venezuela. I use satellite-recorded night light emissions and gridded population data to estimate rural poverty rates at the state and municipality levels. I follow a technical procedure to extract, calibrate, and harmonize nighttime light data from two instruments to align them to population data on a cell area of 1x1 kilometers, covering 2000-2020. The results identify Amazonas, Apure, Barinas, Delta Amacuro, and Guarico as the five states with the highest rural poverty rate in 2020, while the Distrito Capital, Carabobo, Miranda, Nueva Esparta, and Vargas have the lowest rates. I also examine if there has been a significant change in rural poverty throughout the recent economic collapse experienced in Venezuela. In this chapter, I verify how most of the Venezuelan territory had a considerable increase in rural poverty rates since 2013–2014.

The general assumption in this literature is that night light emissions implicitly capture information about economic activity. Elvidge et al. (1997), Sutton and Costanza

[1] For details, see World Bank (2018)'s report on poverty and shared prosperity. Multidimensional poverty is a measure that considers income and access to basic infrastructure, education, health, and security.

[2] In 2020, 91 percent of the world population had access, which is significantly greater than the 70 percent in 1990 and the 83 percent in 2010 (International Energy Agency et al. 2022).

(2002), Doll et al. (2006), Ghosh et al. (2010), Henderson et al. (2012), confirm the relationship between economic activity and lights, the last one proposing a statistical framework to measure actual economic growth. More recently, Bruederle and Hodler (2018) use nighttime lights to estimate human development at a local level, Ferreira (2018) tracks regional socio-economic outcomes for Namibia, and Arderne et al. (2020) identify electrification targets using open data and nighttime lights imagery.

In this context, oil-dependent countries as Venezuela (oil exports accounted for around 75 percent of its total export revenue on average during the last decade and more than 60 percent by 2020), where government efforts of diversification, industrialization, and poverty alleviation proceed from a stable oil market, are less documented. These economies are usually more subject to external shocks (for example, sudden changes in international commodity prices), which might endanger their finances, impacting public policies and the household behavior to migrate or settle down in a specific area. Furthermore, countries with an important oil industry have specificities that may influence how the institutional framework, policy planning, and political economy evolve (Manzano et al. 2008) and do not necessarily invest more in local electrical infrastructure (Min 2010).

This chapter relies on the idea that nighttime lights and spatially distributed population data allow measuring rural poverty, considering any inhabitant living in unlit cells as rural poor (Smith and Wills 2018). This assumption leads to two immediate limitations. First, access to electricity is the only dimension to consider in poverty identification. Second, in this case, I neither identify urban poverty nor cannot rule out its presence. For example, Petare is a neighborhood in the Sucre municipality in the steep hills of northwestern Miranda state, Venezuela. According to 2011 census results, the population in Petare is 372,470. It is still considered one of the largest urban slums in Latin America, with significant access to informal electricity service; therefore, its poverty rate is not accounted for by this approach. Nevertheless, both limitations still allow us to understand any eventual estimate as a lower bound for rural and overall poverty in Venezuela.

Smith and Wills (2018) argue that using darkness to estimate poverty is simpler than identifying poverty in the light and that electrification could significantly explain rural poverty. Electrification is inferior in rural areas, mainly because they are located far from

national grids or urban centers, involving high infrastructure costs to extend electrical grids. Therefore, having access to electricity could signaling policy efforts to alleviate poverty. It should also give a significant marginal return to rural residents, reducing the time allocated to fuelwood collection and increasing labor supply, schooling, household per capita income, and expenditure (Khandker et al. 2014). Furthermore, Smith and Wills (2018) confirm that people tend to switch from kerosene to electric lighting soon after leaving extreme poverty, making access to electricity a good benchmark for rural poverty.

Venezuela is experiencing an economic collapse described in Kurmanaev (2019) as "*...the worst outside of war in Decades*", which impacts poverty indicators. The National Survey of Living Conditions (ENCOVI) found that about 65 percent of Venezuelan households suffered multidimensional poverty in 2021 (ENCOVI 2021).[3] That figure rises to 95 percent when measured by the poverty line, with almost 77 percent in extreme conditions. From 2013 to 2021, the GDP shrank about 75 percent, and the significant impacts of oil price shocks and domestic mismanagement led to a socio-economic crisis and a political situation with no precedent in Latin America (Cerra 2016; Halff et al. 2017; Vera 2017; Olivo and Saboin 2020; Maldonado and Olivo 2022). In 2016, the country entered hyperinflation and from then to the end of 2021, the Central Bank of Venezuela (BCV) reports an inflation rate near 30 billion percent. In this context, Venezuela has been experiencing a massive displacement crisis (more than 7 million Venezuelans have fled the country by September 2022 according to the Inter-Agency Coordination Platform for Refugees and Migrants from Venezuela 2022).

The Venezuelan electrical system has gone into a dramatic deterioration. Electricity rationing was implemented by zones since 2008, and the regime declared an electrical emergency in 2009, announcing generation capacity development projects. Nonetheless, those development projects were not able to provide enough power supply to even support an already constrained demand in part due to a deprofessionalization of the sector and a narrow political criterion imposed on the planning and execution process (González Oquendo 2019; Guevara Baro 2020). Thus, service interruptions have been increasing,

[3]The ENCOVI is a project launched in 2014 and conducted by academic researchers from three top universities of Venezuela: Andrés Bello Catholic University (UCAB), Central University of Venezuela (UCV), and Simón Bolívar University (USB), as a response to the lack of household's microdata collected by the National Institute of Statistics of Venezuela.

with progressively worse blackouts since 2010 (one of the mains leaving in the dark 90 percent of Venezuelan territory for several days in 2019, having adverse humanitarian consequences and sectorial effects as indicated by Sabatini and Patterson 2021). In this sense, the country's energy crisis is still ongoing, and rebuilding the electricity sector is mandatory nowadays (Sabatini and Patterson 2021). For example, the results of the ENCOVI (2019-2020) suggest that 90 percent of households in Venezuela report interruption of electric service, of which 32 percent reported daily failures. This scenario reminds us that it is not only the case that the central government must effectively provide electricity grids, but even with electrical grids installed, there could be long periods of low intensity or lack of lights. This remarks how access to public services, such as electricity, might be a vital predictor of poverty in Venezuela.

Alongside the socio-economic crisis occurs a political turmoil characterized by high polarization and citizen protests. The central claims lie in inferior quality and access to fuel, electricity, and water supply, as well as intense scenarios looking to regain social and civil rights and asking for a government change in 2014, 2017, and 2019. In particular, the regime response was repression by security forces and pro-government armed civilians (Denis 2021; Venezuelan Observatory of Social Conflict 2020). The current Venezuelan context is ideal for shedding light on how poverty is evolving in Venezuela and generating data to support eventual recovery efforts and poverty alleviation programs. This chapter presents an empirical exercise to spatially characterize rural poverty spanning 2000–2020. One main finding confirms how 2013–2014 were critical years from which rural poverty started rising in Venezuela.

The contribution of this chapter is threefold. Firstly, Venezuela has no recent estimations of rural poverty at a geographically disaggregated level, and household microdata do not characterize poverty by urban-rural classification. Is rural poverty the same across all states and municipalities? This chapter fills this gap by estimating rural poverty at subnational levels, matching satellite measures of nighttime lights with a gridded population count as a novel, less-costly alternative to obtain detailed independent results from methodologies based on surveys or administrative data. Spatial techniques are beneficial in reducing measurement errors in existing administrative data, especially in territories with informal settlements. A first approximation to

understanding Venezuela's local dynamics requires recognizing these specificities.[4] A better understanding of the spatial distribution of rural poverty across the Venezuelan territory would allow policymakers to account for geographical locations while setting national and subnational policies.

Secondly, unlike Smith and Wills (2018), I calibrate and harmonize light datasets to generate time series spanning 2000–2020. I also derive the sigmoid relationship suggested by Zhao et al. (2019) and Li et al. (2020) between the light inputs. The sigmoid-like curve reveals a light intensity threshold that separates unlit areas or cells with low light intensity from brighter cells. Every pixel with population and light values below that "rural poverty threshold" defines a rural poor area.

Thirdly, this chapter uses two multitemporal gridded population datasets to estimate rural poverty as an attempt to evaluate robustness and precision. In this sense, it is possible to verify which population distribution yields more precise and consistent estimates over time for Venezuela.

Two main strands of literature could benefit from this work. On the one hand, studies closely related to multidimensional and energy poverty in developing economies. Spagnoletti and O'Callaghan (2013), Khandker et al. (2014), González-Eguino (2015), Lee et al. (2016), Njiru and Letema (2018), Mendoza Jr et al. (2019), Getie (2020) are some examples of studying the role of the access to electricity to poverty alleviation. On the other hand, studies characterizing poverty within regional dynamics and in resource-based countries as in Bazilian et al. (2013), Loayza and Rigolini (2016), and Smith and Wills (2018).

The chapter is organized as follows. Section 3.2 describes the study area and night light data's pre-processing (inter-annual calibration and harmonization). Section 3.3 exposes the methodology and Section 3.4 stylized facts and rural poverty rates results within and across Venezuela. Section 3.5 concludes.

[4]This is not a new challenge. However, the studies highlighting the relevance of understanding rural poverty in Venezuela are from more than two decades ago (Llambí et al. 1994; Márquez 1994; Martel 1995; Riutort 1999).

3.2 Study Area and Data

Venezuela is geographically located in northern South America. It is divided into one Distrito Capital, 15 non-oil producing states (Amazonas, Aragua, Bolívar, Carabobo, Cojedes, Lara, Miranda, Merida, Nueva Esparta, Portuguesa, Sucre, Tachira, Trujillo, Vargas, and Yaracuy), eight oil-producing states (Anzoategui, Apure, Barinas, Delta Amacuro, Falcon, Guarico, Monagas, and Zulia), and Federal dependencies consisting in small offshore islands in the Caribbean Sea and the Gulf of Venezuela. In this exercise, the Federal dependencies are excluded due to the small number of observations.

The characterization of rural poverty at the municipality level covers the period 2000–2020. To construct this sample and obtain estimates, I extract and process gridded datasets of population counts as well as satellite-recorded nighttime lights.

The population data come from two sources. The Oak Ridge National Laboratory, ORNL (2021, March), provides the LandScan Global Population database. It represents an ambient population (average over 24 hours) distribution and is currently available annually from 2000 to 2019. I also use annual top-down unconstrained geospatial population distribution from the repository of WorldPop (2021) spanning the period 2000–2020. ORNL and WorldPop's imagery could be found at 30 arc-second resolution or a cell area of 1x1 kilometers. In both cases, population counts are recalculated for the last five years to consider the severe international migration outflow experienced in Venezuela.[5]

Two instruments capture the light data. On the one hand, I have access to annual satellite imagery of the U.S. Air Force Defense Meteorological Satellite Program (DMSP)-Operational Linescan System for the period 1992–2013. DMSP datasets exist for 30 satellite-years.[6] On the other hand, the Earth Observation Group of the Payne Institute for Public Policy (Colorado School of Mines) produces a consistently processed time series -version 2- of annual global nighttime light imagery from the Visible Infrared Imaging Radiometer Suite (VIIRS) on-board the Suomi National Polar-Orbiting Partnership

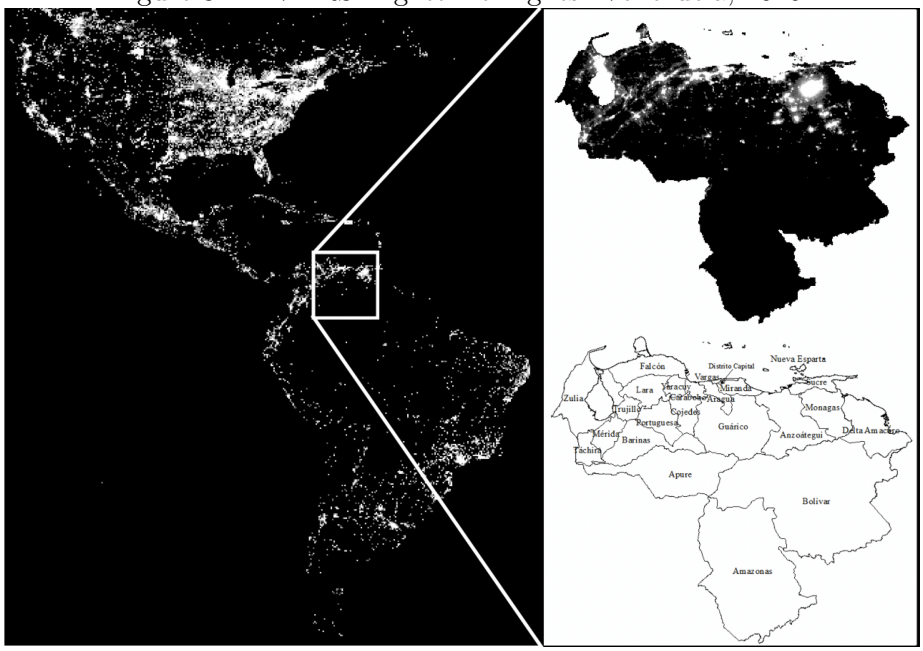
[5]For this, I assume an equally proportionally distributed effect throughout all pixels. The ORNL's files are at https://landscan.ornl.gov, while the WorldPop's files can be found at https://hub.worldpop.org/project/categories?id=3.

[6]There are data from satellite F10 and F12 for the year 1994, F12 and F14 for 1997–1999, F14 and F15 for 2000–2003, and F15 and F16 for 2004–2007. In those years, I use the data from the newer satellite.

satellite platform for the period 2012–2020. Looking for pre-filtered light data, I use the DMSP annual composites of stable lights and the VIIRS annual masked average radiance products (Elvidge et al. 2021).[7]

Figure 3.1 shows the spatial location of Venezuela with its states and the distribution of VIIRS lights across the territory for 2020. Each satellite imagery contains light intensity on a grid, where whiter pixels mean more light intensity and black pixels unlit areas. In the DMSP data, the unit of intensity of light is a six-bit digital number (DN), recorded for every 30 arc-second pixels. The values range from zero (unlit) to 63 (brightest light), and the sensors have no onboard calibration. Unlike DMSP data, the VIIRS raw data has a spatial resolution of 15 arc-seconds, providing more details; it has onboard calibration and a wider radiance range leading to more robust low light detection, thus not suffering from frequent saturation as in DMSP imagery.[8]

Figure 3.1: VIIRS Nighttime Lights: Venezuela, 2020



Source: Earth Observation Group, Payne Institute - Colorado School of Mines.

[7]The files are at the online repository of the Earth Observation Group (2022, June): https://payneinstitute.mines.edu/eog/nighttime-lights/. For details, see the readme file associate with the respective sources.

[8]The measure of light intensity from DMSP data is top-coded at 63 and suffers saturation (Bluhm and Krause 2018). However, this chapter will characterize rural poverty through unlit areas or low-intensity lights, making irrelevant a top-coding correction approach.

It is possible to align the population count pixels with nighttime lights on the same coordinate system, reaching a final sample spanning 2000–2020. Nevertheless, as in Chapter 2, that requires performing extra technical work over the light data. DMSP data has no onboard calibration, so I must carry out an inter-annual calibration procedure from 2000 to 2013. There are also differences between DMSP and VIIRS datasets related to sensor variations in spatial resolution, spectral response, point of spread function, overpass time at night, and wider radiance range of the VIIRS (Elvidge et al. 2013; Li et al. 2017; Zhao et al. 2019; Li et al. 2020). Therefore, the analysis requires a harmonization procedure. In this case, the harmonization would result in identifying a light intensity threshold separating unlit areas or pixels recording very low light intensity from the rest.

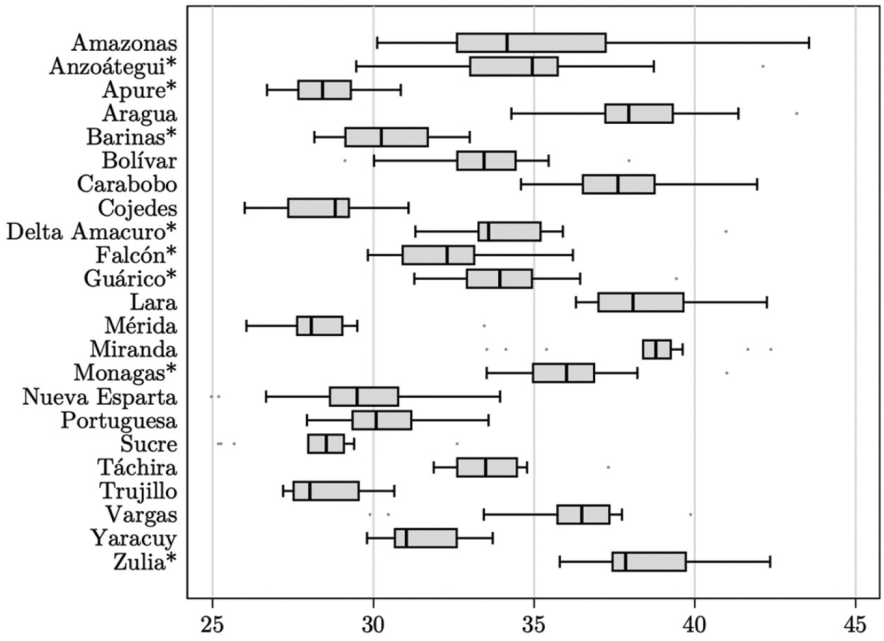
3.2.1 Pre-Processing: Inter-Annual Calibration

Elvidge et al. (2009) propose a general and flexible framework for calibrating DMSP nighttime lights. The idea is to select a region with a relatively stable intensity of lights throughout the entire sample, an area more consistent and invariant across the years in terms of night light emissions. Then, they suggest selecting that region and the satellite-year with its maximum intensity of light as a benchmark or reference to apply a second-order regression model year by year (Equation 3.1). Finally, the calibration consists of using the yearly estimated coefficients to obtain fitted and rounded DN values for all the sample areas ($DN_{adjusted}$).

$$DN_{adjusted} = \hat{\alpha}_0 + \hat{\alpha}_1 DN + \hat{\alpha}_2 DN^2 \tag{3.1}$$

In this case, I calculate the relative mean deviation (RMD) as a measure of the stability of lights. I consider all lit pixels and all the states of Venezuela between 2000 and 2013. Figure 3.2 shows the results in boxplots. Trujillo has a lower median of RMD than the rest of the states, with most of its pixels denoting less relative deviation in the sample. In 2013, the intensity of lights in Trujillo averaged its highest value (12.3). For these reasons, Trujillo is selected as the reference state and 2013 as the base year to apply Equation 3.1. The estimated coefficients from the second-order regression model are in Table 3.1.

Figure 3.2: Boxplots of Relative Mean Deviation, 2000–2013



Source: Own calculations. Note: The boxplots include outliers. Distrito Capital is not considered because of its relatively small number of pixels with respect to the other states. Pixel values of zero are excluded to calculate the relative mean deviation index. *Oil-producing states.

Table 3.1: Coefficients for the Inter-Annual Calibration of DMSP Using Equation 3.1

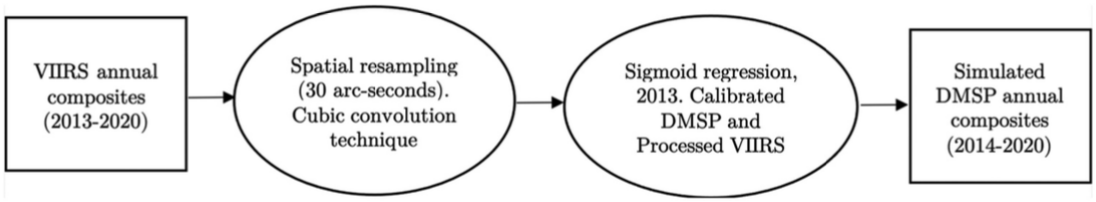
Year	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	R^2
2000	0.4902	0.6503	0.0021	0.8922
2001	0.4896	0.6420	0.0023	0.8931
2002	0.6321	0.7028	0.0014	0.8878
2003	0.6654	0.3426	0.0057	0.8828
2004	1.0809	0.5769	0.0031	0.8962
2005	0.7819	0.4270	0.0054	0.9475
2006	0.7598	0.5679	0.0042	0.9268
2007	1.0371	0.6188	0.0034	0.9290
2008	0.4314	0.6501	0.0033	0.9597
2009	0.4215	0.5717	0.0041	0.9554
2010	0.8457	0.8495	0.0012	0.9527
2011	0.5562	0.6492	0.0044	0.9589
2012	0.3154	0.8131	0.0027	0.9741
2013	0.0000	1.0000	0.0000	1.0000

Note: Trujillo is the reference because of its low relative mean deviation with no outliers in 2000–2013. The maximum intensity of light was reached in 2013 (base satellite-year to calibrate). Total observations per year are 9,579.

3.2.2 Pre-Processing: Harmonization

The idea is to estimate 2014–2020 DMSP annual composites using 2013–2020 VIIRS annual composites as an input. Figure 3.3 shows each step of this procedure.

Figure 3.3: Simulating DMSP Annual Composites, 2014–2020



Source: Own pre-processing.

Firstly, I apply a cubic convolution resampling technique suitable for continuous data, smoothing and scaling the VIIRS spatial resolution to 30 arc-seconds (the same as for DMSP data). Then, I describe the relationship between the values of the calibrated DMSP and the logarithmic transformation of the VIIRS data -adding one to account for unlit pixels- through a sigmoid function with four parameters: a, b, c, and d (Equation 3.2).

$$DN_{simulated} = a + b\left(\frac{1}{1 + e^{-c(Log(\text{VIIRS}+1)-d)}}\right) \tag{3.2}$$

Zhao et al. (2019) and Li et al. (2020) propose this method noting that the radiance variation of processed VIIRS data differs across types of areas. They confirm that a sigmoid function captures adequately that relationship under different lit environments among rural, rural-urban transition zones, and urban cores.

Table 3.2 shows the parameters obtained through the regression. Both datasets have 2012 and 2013 as common years, but 2012 VIIRS data is annualized using the months from April to December. Therefore, to avoid seasonality issues, this chapter derives the parameters only accounting for the relationship in 2013.

Table 3.2: Parameters Obtained through the Sigmoid Regression, 2013

a	b	c	d	R^2	Observations
7.4384	55.3555	1.1929	1.6425	0.7218	164,826

Note: the parameters represent the minimum value from DMSP on the fitting curve (a), the difference between maximum and minimum values (b), maximum slope (c) and processed VIIRS value on the maximum slope (d).

3.3 Methodology

Rural areas relate to less illuminated space than urban areas, which could explain the low or moderate intensity of lights in regions with vast land (Zhao et al. 2019; Li et al. 2020). Agriculture and population dispersion play an important role in rural areas, and agricultural activity emits marginal lights (Keola et al. 2015). In fact, Keola et al. 2015 conclude that agricultural activity is better estimated through land cover datasets rather than using lights independently of the degree of development of a country. In this sense, areas moderately lit may capture most of the information to describe rural features while unlit areas or areas with very low light intensity could relate to rural poverty.

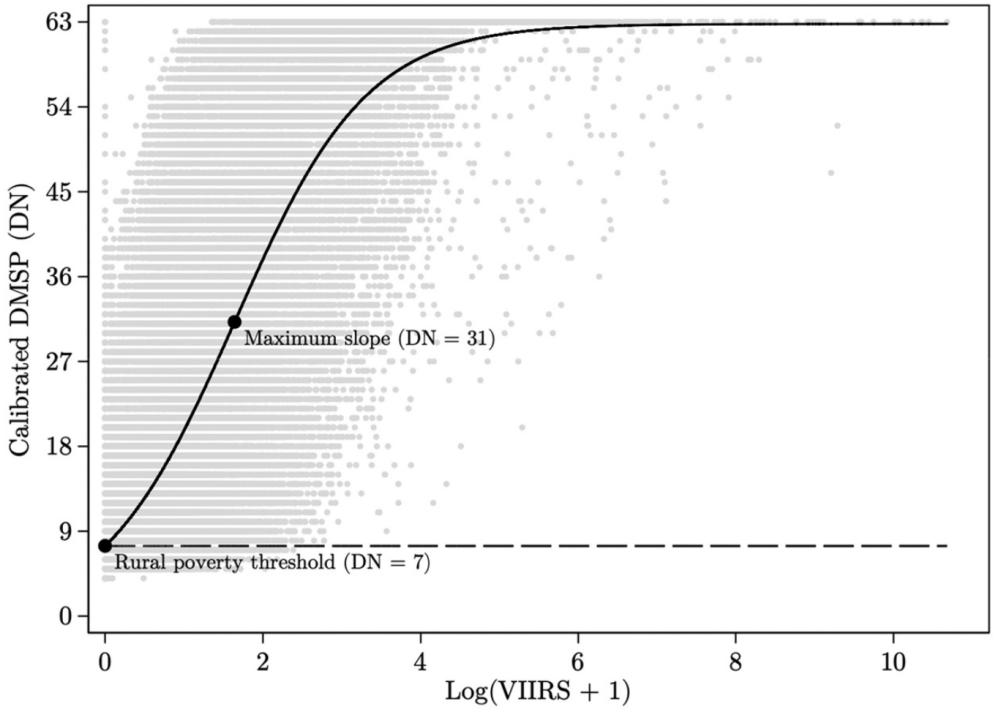
Smith and Wills (2018) suggest that darkness reveals poverty more easily than light. They characterize rural poverty assuming any people living in unlit cells are rural poor. This chapter considers that not only unlit cells but pixels with low intensity of light describe rural poverty. As Sutton (2003) argued, certain human activities always grow slightly around their light source due to blooming effects; thus, most of the low light intensity values from the calibrated and simulated DMSP are not necessarily urban areas. If that is the case, blowing effects might be shading rural features.

In this context, how low should the light intensity be to identify rural poverty? What is the light threshold to define rural poverty? This chapter answers these questions for Venezuela through the technical identification of the sigmoid function proposed by Zhao et al. (2019) and Li et al. (2020). Both studies focus on using estimated parameters to describe human activity in urban cores. However, this methodology also identifies different lit environments allowing to obtain a rural poverty threshold.

Figure 3.4 presents the estimated sigmoid curve from the harmonization procedure. The rural threshold is the minimum value on the fitting curve ($DN = 7$). Values of DMSP greater than 7 are already within the urban-rural transition zone. At the beginning, DN values increase more than proportional to $Log(\text{VIIRS} + 1)$, until the point of maximum slope ($DN = 31$). In this zone, rural areas are transitioning to urban areas. After that point, the DMSP values still increase but less than proportional, tending to urban core values ($DN \to 63$). In this context, the chapter assumes that values in the urban-rural transition zone do not represent rural poverty. Therefore, the identification strategy to define rural poverty is: $DN \leq 7$ with at least one inhabitant

on a cell area of 1x1 kilometers.

Figure 3.4: Relationship between DMSP and VIIRS: Sigmoid Function, 2013



Source: Own calculations. Note: 2013 is a common year with data of lights from DMSP and VIIRS. The sigmoid function reaches its minimum at $7.4384 \approx 7$ DN (threshold to group unlit pixels or pixels with low-intensity values of lights during the harmonization process).

The estimates of rural poverty rates are based on two multitemporal gridded population datasets, ORNL's LandScan and WorldPop, as an attempt to evaluate robustness and precision. Nevertheless, Bustos et al. (2020) already warn us about dissimilarities in gridded population datasets and the relevance of choosing one over another depending on which is best suited for particular research needs in regional analyses.

To reduce volatility and noises regarding measurement errors or satellite sensitivity not captured during the calibration procedure, I create a measure of agglomeration grouping every grid, fulfilling the decision criterion for rural poverty. This allows aggregating the rural poverty rate estimates at the municipality level, weighting by the number of pixels.

This chapter has two evident limitations. On the one hand, the identification of rural poverty is based on access to electricity, which is only one category of multi-dimensional

poverty. On the other hand, I neither identify urban poverty nor cannot rule out its presence. Nevertheless, both limitations allow us to understand any eventual estimate from this approach as a lower bound for rural and overall poverty.

3.4 Results

Table 3.3 and Table 3.4 summarize statistics for Venezuela and its states, respectively. On average, around 93 percent of the Venezuelan territory has a light intensity below the rural poverty threshold. WorldPop data show more populated territory than ORNL data. According to WorldPop, about 55.6 percent of total pixels have at least one inhabitant compared to 43.2 percent when using ORNL. The population density is similar independently of the source, but the population count reveals a major difference (74,666 with ORNL versus 13,294 with WorldPop); those values correspond to the same municipality: Sucre, Miranda. More pixels with WorldPop data satisfy the decision criterion for rural poverty (42.8 percent of total pixels).

Table 3.3: Summary Statistics (Pixel-Based), 2000–2020

	Mean	Standard deviation	Minimum	Maximum	Observations
States: All					
Intensity of light $\leq 7DN$	0.9256	0.2624	0	1	22,608,096
Population, ORNL	0.4324	0.4954	0	1	21,531,520
Rural poverty, ORNL	0.3703	0.4829	0	1	21,531,520
Population count, ORNL	25.7496	482.7273	0	74,666	21,531,520
Population, WorldPop	0.5560	0.4969	0	1	22,608,096
Rural poverty, WorldPop	0.4823	0.4997	0	1	22,608,096
Population count, WorldPop	25.7800	214.1708	0	13,294	22,608,096
States: Non-oil					
Intensity of light $\leq 7DN$	0.9416	0.2346	0	1	13,539,540
Population, ORNL	0.2722	0.4451	0	1	12,894,800
Rural poverty, ORNL	0.2216	0.4153	0	1	12,894,800
Population count, ORNL	27.9680	548.5228	0	74,666	12,894,800
Population, WorldPop	0.3687	0.4825	0	1	13,539,540
Rural poverty, WorldPop	0.3110	0.4629	0	1	13,539,540
Population count, WorldPop	28.7864	234.7769	0	13,294	13,539,540
States: Oil					
Intensity of light $\leq 7DN$	0.9018	0.2976	0	1	9,068,556
Population, ORNL	0.6718	0.4696	0	1	8,636,720
Rural poverty, ORNL	0.5924	0.4914	0	1	8,636,720
Population count, ORNL	22.4375	362.9097	0	54,111	8,636,720
Population, WorldPop	0.8357	0.3706	0	1	9,068,556
Rural poverty, WorldPop	0.7381	0.4396	0	1	9,068,556
Population count, WorldPop	21.2915	178.9511	0	8868	9,068,556

Note: When writing this chapter, ORNL population data is available until 2019.

Table 3.4: Summary of Nighttime Lights and Population by State, 2000–2020

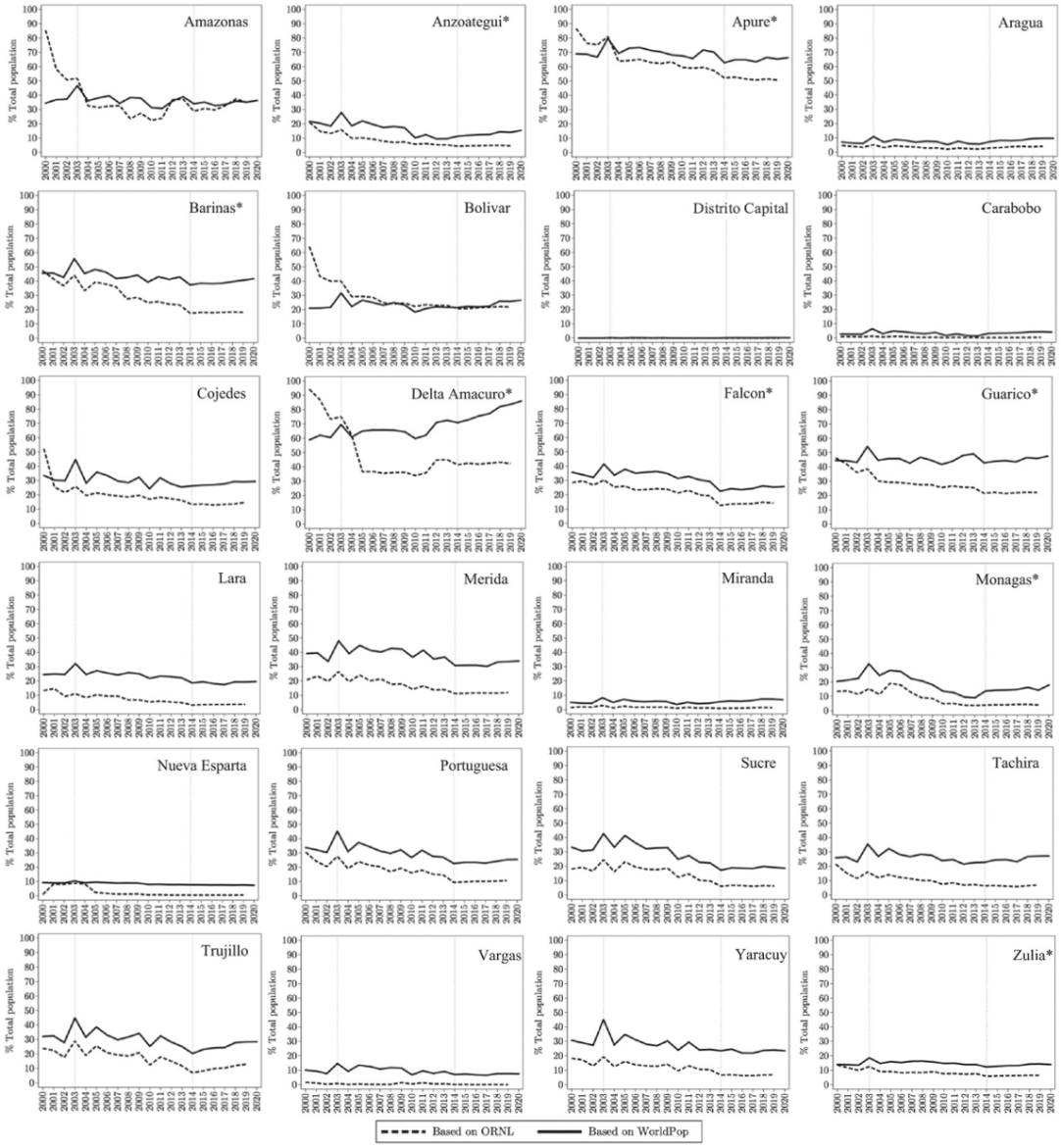
	Intensity of light (DN)				Population, ORNL		Population, WorldPop	
	0 - 7	8 - 30	31 - 63	Gini index	At least one inhabitant (total area)	Average density per 1 km^2	At least one inhabitant (total area)	Average density per 1 km^2
States: Non-oil								
Amazonas	99.9%	0.1%	0.0%	0.3794	4.2%	0.5	4.1%	0.5
Bolivar	98.6%	1.2%	0.2%	0.3550	19.5%	2.3	29.4%	4.8
Cojedes	90.1%	9.0%	1.0%	0.3249	80.2%	17.7	95.1%	19.0
Merida	86.6%	12.1%	1.3%	0.3134	66.7%	58.4	100.0%	56.3
Portuguesa	84.9%	13.6%	1.6%	0.3344	78.9%	39.6	98.4%	38.8
Lara	84.3%	12.8%	2.9%	0.3810	68.8%	84.5	100.0%	81.2
Tachira	80.9%	16.2%	3.0%	0.3489	80.7%	94.4	100.0%	95.7
Sucre	78.7%	19.3%	2.0%	0.3234	80.3%	71.8	99.6%	69.3
Yaracuy	74.4%	22.6%	3.0%	0.3183	76.3%	74.6	100.0%	83.8
Trujillo	72.4%	24.9%	2.7%	0.3124	73.2%	74.5	100.0%	74.3
Aragua	66.2%	25.8%	8.0%	0.3820	81.7%	165.7	99.5%	207.5
Miranda	57.8%	29.6%	12.6%	0.3685	73.6%	259.2	99.9%	275.1
Vargas	49.0%	39.4%	11.5%	0.3500	43.6%	281.2	100.0%	300.6
Carabobo	33.4%	43.3%	23.3%	0.3749	69.9%	455.6	91.4%	410.1
Nueva Esparta	27.8%	38.5%	33.7%	0.3134	91.7%	364.5	100.0%	393.0
Distrito Capital	1.3%	18.2%	80.5%	0.1747	93.4%	6,282.1	100.0%	5189.6
States: Oil								
Delta Amacuro	98.8%	1.0%	0.2%	0.3485	38.4%	2.7	65.7%	3.6
Apure	98.6%	1.3%	0.1%	0.2934	72.2%	3.7	66.6%	4.9
Guarico	95.1%	4.3%	0.6%	0.3529	82.0%	11.5	90.8%	9.1
Barinas	93.2%	6.1%	0.7%	0.3311	74.1%	18.6	95.7%	19.1
Falcon	88.3%	10.3%	1.4%	0.3513	79.3%	27.7	98.4%	28.0
Zulia	84.9%	11.9%	3.2%	0.3825	54.3%	68.5	98.1%	71.6
Anzoategui	79.0%	17.5%	3.6%	0.3654	72.0%	33.4	75.0%	26.4
Monagas	70.8%	21.2%	8.0%	0.3720	51.4%	34.4	97.2%	26.6

Note: States sorted by proportion of zero/very low intensity of lights, DN 0–7. Pixel values with DN 0–7 are excluded to calculate the Gini index. When writing this chapter, ORNL population data is available until 2019.

The eight oil-producing states cover about 40 percent of the total territory, but their populated pixels more than double those of non-oil states. The light distribution does not reveal this significant difference. Five out of the 23 states plus the Distrito Capital have less than 70 percent of their area below the rural poverty threshold. Neither of those states produces oil. Furthermore, excluding unlit areas and the Distrito Capital, lights seem to be similarly sparse within each state. The top states revealing significant gaps between the extension of lights and population are Amazonas, Bolivar, Carabobo, Delta Amacuro, Nueva Esparta, and the Distrito Capital. To a lesser extent, Monagas and Zulia. In all those cases, only Carabobo, Nueva Esparta, and the Distrito Capital consistently have more areas with zero or low light intensity than populated areas.

The results of rural poverty as a percentage of the state-level population are in Figure 3.5.

Figure 3.5: Rural Poverty Rate by State



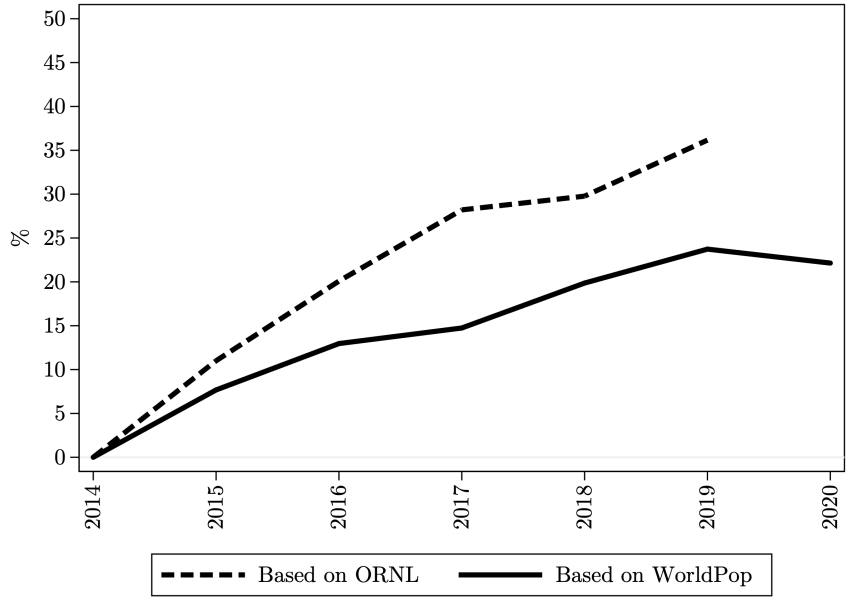
Source: Own calculations. Note: *Oil-producing state.

In general, the satellite sensitivity reveals volatility during 2003–2004 and, in most cases, an absolute peak in 2003. During the period 2002–2003, Venezuela experienced significant political turmoil, leading to a general strike and the onset of exchange rate controls. There are also broad trends. For example, the years 2013–2014 seem to reveal

a common "before and after". Before that period, there was general stability or even a slight downward trend (i.e., mostly between 2009 and 2013). Although the Venezuelan GDP shrank in 2009 and 2010, in 2010, there were parliamentary elections, and in 2012, there were presidential elections. As of 2013–2014, which coincides with the beginning of the current recession, rural poverty generally increased.

I find evidence that 2013–2014 was a switching period for rural areas in Venezuela using both sources, from which rural poverty has been spatially "expanding". I calculate the number of populated pixels with values below the rural poverty threshold to calculate the relative change to 2014 (base year). Figure 3.6 shows how the maximum values of the relative change reduced over time until 2014. All those relative differences are positive after 2014, confirming that unlit pixels with population started gaining extension in the presence of the ongoing economic collapse in Venezuela.

Figure 3.6: Rural Poverty Agglomeration Switch (Base Year 2014)

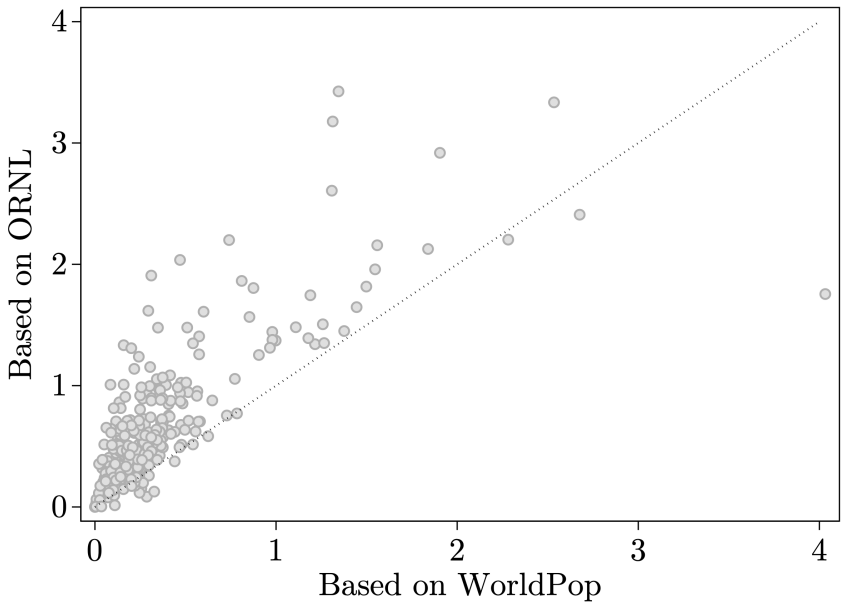


Source: Own calculations. Note: The values represent the maximum relative change of each year to 2014 over the full sample at the state level.

Rural poverty rate estimates with WorldPop are more consistent over time, and they are usually above the results from ORNL. Both ORNL and WorldPop have consistently shown higher accuracy in approximating the known population (Bustos et al. 2020).

However, Bustos et al. (2020) specifically suggest using WorldPop over ORNL for determining populated areas and population estimates along with thresholds in low-density areas (i.e., rural regions). Furthermore, Figure 3.7 represents the relative dispersion of each municipality result around its mean by population dataset throughout 2000–2020 (i.e., the coefficient of variation or the ratio of the standard deviation to the mean). In this case, results with ORNL show higher values compared to WorldPop, indicating more precise estimates from the latter. Therefore, even if both sources led to similar broad trends, this chapter relies mostly on the estimates with WorldPop.

Figure 3.7: Coefficient of Variation by Municipality, 2000-2020



Source: Own calculations. Note: The bisector is used as a reference.

Table 3.5 reports summary statistics and estimates of rural poverty rates for 2014 and 2020 based on WorldPop at the state and municipality level. The top five states with the highest rural poverty rate in 2020 are Amazonas (90 percent), Delta Amacuro (84.8 percent), Apure (71.2 percent), Barinas (60.6 percent), and Guarico (57.6 percent), while the bottom five are Nueva Esparta (11.9 percent), Miranda (10 percent), Vargas (7.6 percent), Carabobo (7.0 percent), and the Distrito Capital (0.1 percent).

Table 3.5: Rural Poverty Rate Based On WorldPop by Municipality

State and Municipality	Mean	Standard deviation	Minimum	Maximum	2014	2020
Amazonas	**85.4%**	**8.5%**	**61.8%**	**91.9%**	**76.1%**	**90.0%**
Alto Orinoco	100.0%	0.0%	99.8%	100.0%	100.0%	100.0%
Atabapo	96.7%	3.8%	88.9%	100.0%	90.0%	100.0%
Manapiare	84.1%	24.1%	28.6%	100.0%	28.6%	100.0%
Maroa	90.3%	16.6%	35.6%	100.0%	87.8%	100.0%
Rio Negro	95.7%	10.5%	55.6%	100.0%	100.0%	100.0%
Autana	99.6%	0.8%	97.7%	100.0%	98.0%	99.5%
Atures	31.3%	3.9%	26.4%	43.4%	28.2%	30.2%
Delta Amacuro*	**76.3%**	**6.6%**	**62.6%**	**85.9%**	**78.7%**	**84.8%**
Antonio Diaz	99.7%	0.3%	99.0%	100.0%	99.0%	100.0%
Pedernales	72.9%	18.0%	33.7%	96.7%	87.6%	96.7%
Casacoima	71.8%	5.0%	62.5%	79.3%	68.9%	79.3%
Tucupita	60.9%	3.1%	55.1%	67.7%	59.3%	63.3%
Apure*	**73.4%**	**5.5%**	**64.5%**	**84%**	**67.5%**	**71.2%**
Pedro Camejo	85.9%	6.0%	75.9%	98.8%	81.7%	85.5%
Achaguas	83.4%	3.7%	77.6%	91.0%	77.6%	82.7%
Muñoz	81.2%	5.6%	73.0%	91.9%	73.0%	76.8%
Romulo Gallegos	71.7%	3.2%	63.7%	75.6%	69.9%	74.9%
Paez	61.7%	3.5%	54.8%	68.7%	63.0%	65.5%
San Fernando	65.2%	5.4%	58.8%	80.2%	58.8%	61.2%
Biruaca	65.1%	11.0%	47.7%	81.7%	48.3%	52.1%
Barinas*	**62.7%**	**6.4%**	**52.7%**	**76.3%**	**55.8%**	**60.6%**
Andres Eloy Blanco	89.9%	3.6%	84.0%	96.2%	84.0%	89.8%
Cruz Paredes	80.7%	5.7%	64.3%	89.7%	79.3%	85.7%
Bolivar	73.1%	3.9%	67.1%	79.7%	69.9%	78.3%
Sosa	84.6%	7.2%	75.0%	98.2%	75.0%	76.0%
Pedraza	75.5%	5.3%	67.8%	85.9%	68.0%	74.7%
Arismendi	68.1%	5.3%	56.0%	80.9%	64.0%	70.5%
Ezequiel Zamora	66.0%	2.9%	60.8%	72.3%	63.9%	67.9%
Antonio Jose de Sucre	53.3%	3.7%	47.2%	64.6%	53.4%	54.5%
Rojas	74.3%	17.2%	49.3%	97.0%	49.3%	53.2%
Obispos	45.8%	10.4%	32.4%	73.1%	32.4%	37.3%
Alberto Arvelo Torrealba	25.2%	8.2%	15.2%	50.1%	17.0%	21.7%
Barinas	16.2%	3.3%	12.9%	27.6%	13.0%	17.2%
Guarico*	**54.5%**	**4.8%**	**46.6%**	**65.3%**	**51.7%**	**57.6%**
Santa Maria de Ipire	65.4%	15.7%	34.7%	86.9%	70.0%	86.9%
San Geronimo de Guayabal	80.7%	3.9%	74.4%	91.8%	74.4%	80.6%
San Jose de Guaribe	69.9%	4.3%	65.2%	79.8%	65.3%	71.2%
Ribas	71.6%	4.9%	63.6%	80.5%	63.6%	70.8%
Ortiz	69.9%	5.5%	61.2%	86.7%	63.4%	70.5%
El Socorro	68.6%	4.8%	60.0%	80.8%	66.6%	69.3%
Chaguaramas	63.2%	4.9%	57.5%	76.2%	58.4%	67.9%
Monagas	56.7%	4.6%	50.4%	65.7%	53.4%	60.9%
Las Mercedes	48.8%	5.7%	36.2%	61.5%	49.1%	55.2%
Miranda	46.9%	3.3%	43.1%	57.1%	46.7%	49.8%
Mellado	50.5%	4.1%	44.8%	61.1%	48.1%	48.8%
Zaraza	43.0%	2.5%	37.6%	48.2%	40.4%	46.1%
Infante	41.5%	2.7%	37.2%	48.0%	39.5%	44.6%
Camaguan	21.6%	1.7%	19.1%	26.8%	20.7%	22.3%
Roscio	19.6%	3.1%	14.4%	28.9%	16.5%	18.7%
Bolivar	**55.9%**	**5.1%**	**46.8%**	**65.9%**	**53.9%**	**57.2%**
Sucre	96.0%	1.9%	91.9%	98.7%	95.4%	97.5%
Sifontes	94.2%	2.2%	90.3%	97.5%	94.0%	95.6%
Cedeño	90.8%	2.2%	86.9%	94.6%	89.5%	93.4%
Gran Sabana	51.5%	12.9%	26.4%	70.6%	61.2%	70.6%
Padre Pedro Chien	78.7%	9.6%	62.4%	91.5%	62.4%	70.5%
Angostura	57.7%	5.4%	45.1%	67.2%	60.2%	67.2%
Piar	46.3%	5.8%	37.7%	61.5%	40.8%	43.4%
El Callao	37.3%	5.3%	28.6%	49.2%	32.6%	33.0%
Roscio	31.6%	3.6%	24.5%	43.0%	30.8%	30.8%
Heres	23.6%	3.5%	17.8%	32.7%	22.8%	21.5%
Caroni	7.3%	3.7%	3.2%	18.5%	3.8%	6.0%
Merida	**57.4%**	**8.8%**	**44.3%**	**73.3%**	**45.9%**	**51.5%**
Aricagua	99.5%	0.8%	97.7%	100.0%	99.2%	99.3%
Guaraque	98.5%	2.1%	94.7%	100.0%	94.9%	96.5%
Justo Briceño	97.9%	2.6%	91.7%	100.0%	91.7%	96.5%
Arzobispo Chacon	97.1%	4.2%	90.5%	100.0%	91.7%	92.8%
Andres Bello	72.8%	8.9%	61.0%	92.6%	63.0%	71.3%
Padre Noguera	58.6%	9.1%	46.7%	87.1%	46.7%	68.3%
Caracciolo Parra Olmedo	72.9%	8.5%	60.7%	86.2%	62.7%	65.4%
Julio Cesar Salas	60.6%	10.8%	40.9%	82.8%	40.9%	57.2%
Santos Marquina	77.1%	16.9%	52.2%	97.1%	55.8%	57.2%
Sucre	61.7%	6.5%	51.1%	73.9%	54.3%	56.4%
Antonio Pinto Salinas	60.4%	11.7%	45.8%	80.1%	47.5%	54.6%
Obispo Ramos de Lora	57.8%	12.4%	38.2%	77.4%	39.0%	49.9%
Cardenal Quintero	73.3%	18.4%	47.1%	97.4%	47.5%	47.9%
Zea	40.6%	9.1%	25.0%	58.6%	25.0%	47.0%
Tulio Febres Cordero	46.0%	4.9%	36.9%	55.9%	37.0%	44.0%
Rangel	50.9%	13.5%	32.8%	80.2%	37.4%	35.2%
Miranda	31.3%	7.0%	19.5%	45.9%	21.4%	27.5%
Pueblo Llano	44.7%	21.3%	15.9%	87.8%	15.9%	25.3%
Alberto Adriani	22.3%	4.9%	15.1%	33.4%	19.8%	24.2%
Campo Elias	22.5%	3.1%	18.2%	31.8%	21.2%	21.8%
Libertador	21.8%	2.2%	18.1%	26.3%	20.3%	20.4%
Rivas Davila	31.1%	15.8%	9.2%	56.2%	10.7%	14.2%
Tovar	19.9%	7.5%	10.7%	34.2%	10.8%	10.7%

Note: The results are sorted by 2020 estimates. *Oil-producing state.

Table 3.5: (Continued)

State and Municipality	Mean	Standard deviation	Minimum	Maximum	2014	2020
Falcon*	**57.2%**	**8.3%**	**43.4%**	**70.8%**	**45.0%**	**50.2%**
Democracia	86.5%	4.9%	77.1%	95.9%	77.1%	83.9%
Jacura	92.4%	8.7%	76.3%	100.0%	76.3%	83.3%
Union	85.5%	5.8%	73.8%	97.4%	73.8%	82.3%
Petit	85.3%	9.1%	66.5%	98.9%	66.5%	81.4%
Piritu	89.5%	8.3%	75.3%	100.0%	76.4%	80.8%
Bolivar	81.8%	9.1%	65.4%	95.4%	65.7%	75.3%
Palma Sola	91.6%	12.0%	67.5%	100.0%	85.8%	72.2%
Buchivacoa	81.6%	13.7%	56.9%	95.6%	56.9%	66.9%
Urumaco	73.8%	10.0%	55.2%	89.8%	55.2%	66.0%
Acosta	67.1%	6.8%	55.9%	84.3%	55.9%	61.9%
Mauroa	68.9%	9.9%	49.7%	79.6%	49.7%	60.7%
Sucre	69.9%	16.5%	46.6%	97.4%	46.6%	56.5%
Federacion	65.0%	13.1%	44.9%	83.0%	46.3%	49.9%
Dabajuro	50.3%	4.2%	43.4%	62.8%	44.6%	48.1%
San Francisco	55.9%	9.9%	41.1%	72.6%	43.8%	45.3%
Monseñor Iturriza	37.8%	5.3%	25.9%	50.5%	34.4%	41.0%
Cacique Manaure	41.9%	6.4%	32.9%	54.5%	32.9%	37.8%
Tocopero	46.1%	12.1%	31.0%	74.9%	32.5%	37.7%
Falcon	40.4%	12.0%	23.9%	65.2%	23.9%	32.6%
Zamora	34.3%	7.9%	22.8%	48.5%	24.9%	25.3%
Silva	25.1%	2.7%	20.9%	30.7%	20.9%	22.2%
Colina	30.7%	10.4%	16.0%	49.2%	16.0%	19.5%
Miranda	16.2%	3.1%	11.0%	22.9%	11.0%	14.1%
Carirubana	8.9%	3.3%	4.7%	16.5%	4.8%	5.4%
Los Tanques	3.1%	1.1%	1.0%	5.0%	4.0%	5.0%
Cojedes	**47.6%**	**6.3%**	**38.2%**	**61.8%**	**41.3%**	**44.8%**
Pao de San Juan Bautista	92.9%	6.7%	80.0%	100.0%	80.0%	88.2%
Girardot	82.0%	4.0%	74.0%	88.6%	81.0%	86.4%
Anzoategui	78.7%	13.4%	59.0%	100.0%	60.4%	67.7%
Romulo Gallegos	49.6%	5.0%	43.8%	61.8%	45.9%	45.6%
Ricaurte	47.0%	7.3%	39.1%	70.5%	39.1%	41.8%
Tinaco	33.2%	5.6%	26.4%	47.3%	26.9%	27.6%
Ezequiel Zamora	19.3%	4.7%	12.5%	35.3%	17.9%	20.2%
Falcon	13.5%	5.1%	1.8%	27.3%	12.6%	15.7%
Lima Blanco	12.6%	4.6%	6.9%	25.8%	7.6%	9.5%
Tachira	**42.0%**	**8.1%**	**30.1%**	**57.4%**	**34.5%**	**41.8%**
Francisco de Miranda	99.8%	0.5%	98.1%	100.0%	98.1%	100.0%
Uribante	91.2%	2.7%	86.0%	97.5%	86.0%	89.6%
Panamericano	76.4%	7.8%	64.1%	91.0%	69.7%	72.1%
Samuel Dario Maldonado	73.6%	12.5%	52.0%	93.8%	64.7%	70.9%
Fernandez Feo	69.1%	6.2%	59.7%	79.9%	61.2%	69.2%
Rafael Urdaneta	74.8%	20.2%	38.0%	100.0%	38.0%	67.2%
Libertador	72.1%	9.9%	55.3%	83.9%	55.3%	66.3%
Sucre	74.1%	13.7%	57.3%	100.0%	59.1%	64.3%
Jose Maria Vargas	61.2%	18.3%	33.1%	96.2%	33.1%	56.2%
Garcia de Hevia	55.4%	3.5%	48.3%	62.2%	52.6%	53.8%
Seboruco	48.0%	11.3%	29.9%	73.2%	39.7%	51.6%
Cordoba	48.0%	4.3%	41.0%	59.2%	42.2%	49.2%
Simon Rodriguez	63.7%	21.6%	34.7%	97.5%	34.7%	48.2%
Antonio Romulo Costa	52.8%	14.9%	33.9%	85.4%	41.9%	39.2%
Libertad	13.5%	17.1%	0.0%	45.1%	20%	38.9%
Ayacucho	34.6%	5.8%	24.2%	47.2%	32.2%	37.6%
San Judas Tadeo	39.8%	10.9%	25.8%	64.6%	30.2%	33.8%
Jauregui	39.3%	6.5%	28.1%	50.9%	31.4%	33.5%
Torbes	33.3%	6.7%	22.6%	46.7%	25.0%	29.6%
Michelena	19.3%	8.8%	5.1 %	38.6%	11.2%	28.5%
Lobatera	13.6%	7.4%	3.8%	28.6%	15.8%	26.5%
Andres Bello	12.9%	3.2%	7.6%	18.7%	8.0%	16.8%
Bolivar	11.7%	7.0%	0.9%	31.6%	12.3%	14.4%
Junin	10.2%	3.8%	5.5%	20.8%	8.4%	13.5%
Cardenas	14.3%	2.1%	10.4%	18.5%	13.7%	11.9%
San Cristobal	9.2%	2.2%	6.2%	13.6%	8.2%	11.7%
Independencia	2.5%	3.6%	0.0%	10.0%	3.4%	9.8%
Pedro Maria Ureña	4.8%	2.3%	1.1%	10.4%	3.9%	7.9%
Guasimos	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Portuguesa	**44.8%**	**8.5%**	**33.1%**	**62.7%**	**35.3%**	**39.4%**
Papelon	87.5%	7.3%	75.6%	96.2%	75.6%	82.6%
Monseñor Jose Vicente de Unda	77.4%	8.7%	64.0%	92.6%	65.1%	69.6%
Guanarito	64.7%	4.0%	59.5%	73.2%	62.5%	65.5%
Santa Rosalia	81.1%	15.1%	56.7%	96.1%	56.7%	62.1%
San Genaro de Boconoito	58.0%	13.4%	42.8%	92.2%	43.2%	47.1%
Ospino	54.8%	10.5%	40.6%	72.1%	40.6%	44.4%
Sucre	50.6%	12.6%	32.0%	72.5%	32.4%	43.1%
Turen	52.8%	10.7%	37.7%	76.4%	37.7%	41.7%
Agua Blanca	27.4%	12.9%	9.1%	65.8%	19.3%	26.9%
Guanare	23.1%	5.4%	17.2%	38.3%	17.5%	18.9%
San Rafael de Onoto	13.4%	5.0%	6.8%	27.3%	17.2%	17.7%
Araure	10.7%	3.8%	5.3%	22.8%	10.8%	13.1%
Esteller	15.4%	4.6%	9.4%	28.9%	9.4%	11.6%
Paez	10.4%	4.3%	6.3%	23.6%	6.9%	7.8%

Note: The results are sorted by 2020 estimates. *Oil-producing state.

Table 3.5: (Continued)

State and Municipality	Mean	Standard deviation	Minimum	Maximum	2014	2020
Trujillo	**39.7%**	**8.2%**	**26.4%**	**57.1%**	**28.9%**	**38.0%**
Juan Vicente Campo Elias	98.4%	3.5%	85.0%	100.0%	95.5%	100.0%
Jose Felipe Marquez Cañizalez	90.9%	8.1%	78.3%	100.0%	80.6%	83.2%
Bocono	61.3%	7.2%	46.8%	72.0%	46.8%	62.6%
Monte Carmelo	70.0%	12.3%	47.8%	91.7%	47.8%	61.6%
Andres Bello	64.0%	15.6%	40.2%	96.8%	40.2%	58.4%
Motatan	44.6%	7.9%	28.3%	60.0%	35.9%	51.2%
Carache	55.6%	10.1%	41.4%	75.5%	41.4%	45.4%
Pampan	37.4%	6.9%	24.6%	57.3%	24.6%	38.1%
Miranda	37.5%	10.4%	23.2%	65.7%	24.8%	37.2%
Candelaria	45.6%	15.1%	24.8%	80.9%	24.8%	36.9%
Sucre	32.9%	9.6%	10.9%	51.5%	17.6%	32.9%
Urdaneta	36.1%	12.4%	17.0%	63.0%	17.0%	28.6%
La Ceiba	37.4%	20.8%	13.5%	92.9%	13.5%	25.0%
Trujillo	21.7%	4.3%	15.3%	32.2%	15.3%	22.5%
Rafael Rangel	16.8%	5.8%	9.2%	26.2%	19.5%	21.9%
Bolivar	16.9%	4.2%	9.8%	26.6%	11.5%	18.4%
Pampanito	7.9%	3.3%	2.3%	14.1%	9.2%	14.1%
Escuque	9.5%	2.2%	6.9%	13.0%	8.4%	12.9%
Valera	6.6%	3.1%	3.3%	15.1%	4.0%	6.1%
San Rafael de Carvajal	1.9%	1.5%	0.0%	6.6%	0.5%	2.4%
Lara	**41.7%**	**5.9%**	**33.4%**	**54.8%**	**35.0%**	**37.0%**
Urdaneta	76.2%	8.2%	64.4%	86.2%	65.7%	66.8%
Crespo	63.6%	4.2%	56.9%	74.6%	61.7%	60.7%
Andres Eloy Blanco	58.5%	5.0%	50.4%	71.6%	52.0%	58.1%
Torres	57.4%	7.7%	44.9%	70.3%	46.8%	48.8%
Moron	51.4%	8.0%	38.8%	65.4%	40.0%	44.4%
Simon Planas	37.3%	8.7%	28.0%	62.2%	28.4%	32.4%
Jimenez	20.5%	8.4%	10.1%	41.9%	11.9%	12.7%
Iribarren	9.8%	2.1%	7.0%	16.6%	8.0%	8.4%
Palavecino	0.7%	1.0%	0.0%	4.5%	0.5%	0.8%
Zulia*	**36.2%**	**4.7%**	**29.0%**	**46.4%**	**30.8%**	**33.6%**
Catatumbo	88.3%	2.0%	82.9%	91.9%	86.7%	89.8%
Jesus Maria Semprun	78.8%	3.3%	70.7%	83.9%	78.3%	79.4%
Francisco Javier Pulgar	76.7%	10.7%	59.7%	91.8%	61.5%	68.7%
Almirante Padilla	57.7%	6.5%	45.1%	76.3%	54.7%	63.0%
Colon	66.0%	6.7%	55.6%	78.7%	57.2%	59.9%
Machiques de Perija	54.4%	3.2%	48.8%	60.4%	54.4%	55.8%
Sucre	60.8%	10.4%	44.3%	79.4%	44.5%	50.7%
Baralt	43.8%	9.4%	30.2%	66.3%	30.2%	40.2%
Guajira	54.9%	14.3%	35.0%	77.0%	35.0%	39.2%
Rosario de Perija	43.4%	3.3%	38.6%	52.4%	43.3%	39.0%
Valmore Rodriguez	36.5%	5.5%	29.5%	54.9%	29.5%	33.5%
Miranda	22.2%	4.5%	15.0%	31.9%	15.0%	18.4%
Mara	18.4%	2.6%	14.4%	26.0%	14.4%	16.2%
Jesus Enrique Lossada	14.7%	3.0%	11.4%	23.8%	13.4%	15.4%
La Cañada de Urdaneta	13.8%	2.2%	9.7%	19.2%	9.7%	13.7%
Santa Rita	11.9%	3.7%	6.6%	21.8%	7.9%	9.8%
Lagunillas	9.1%	2.0%	6.2%	14.5%	6.2%	7.1%
Cabimas	6.0%	2.2%	3.3%	11.3%	3.3%	3.6%
Simon Bolivar	3.5%	2.3%	1.6%	11.7%	1.7%	2.6%
Maracaibo	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
San Francisco	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Sucre	**44.4%**	**13.1%**	**27.0%**	**66.8%**	**27.3%**	**30.4%**
Benitez	76.9%	5.4%	67.9%	87.0%	67.9%	71.8%
Libertador	79.6%	18.7%	54.2%	98.3%	54.2%	57.1%
Andres Eloy Blanco	61.5%	14.7%	39.2%	83.1%	39.2%	44.9%
Cruz Salmeron Acosta	57.2%	13.5%	39.1%	82.2%	39.1%	41.6%
Cajigal	54.5%	20.3%	27.5%	93.7%	27.5%	34.4%
Bolivar	47.0%	13.6%	27.6%	77.3%	27.6%	32.4%
Mariño	44.1%	15.8%	24.5%	75.3%	24.5%	28.6%
Montes	41.5%	12.3%	25.9%	60.5%	29.1%	28.0%
Ribero	39.2%	12.8%	22.2%	62.3%	22.2%	26.9%
Valdez	33.4%	8.1%	24.5%	46.0%	24.5%	26.3%
Arismendi	44.5%	16.5%	19.5%	62.7%	19.5%	23.3%
Mejia	50.1%	25.9%	18.7%	93.7%	18.7%	22.8%
Sucre	14.3%	4.3%	10.2%	25.3%	10.3%	10.2%
Andres Mata	18.9%	10.8%	4.7%	42.0%	4.7%	7.4%
Bermudez	2.9%	3.2%	0.2%	13.3%	0.3%	0.4%
Anzoategui*	**30.7%**	**7.6%**	**19.4%**	**46.1%**	**23.7%**	**29.1%**
Monagas	47.3%	15.5%	26.7%	76.5%	47.0%	74.5%
Cajigal	69.9%	9.0%	55.4%	91.2%	55.4%	65.8%
General Sir Arthur McGregor	59.7%	5.2%	50.1%	72.1%	55.0%	62.9%
Libertad	52.5%	5.8%	41.5%	65.6%	43.8%	53.4%
Miranda	50.0%	12.8%	33.4%	76.7%	38.5%	48.8%
Guanipa	44.6%	12.4%	24.5%	67.0%	31.4%	42.2%
San Juan de Capistrano	69.6%	25.9%	34.9%	98.9%	38.7%	39.1%
Aragua	42.0%	5.3%	34.8%	56.1%	34.8%	37.8%
Bruzual	45.4%	8.5%	33.1%	61.3%	33.1%	35.5%
Carvajal	36.4%	7.1%	23.8%	53.3%	32.0%	32.6%
Simon Rodriguez	23.0%	14.4%	0.2%	52.7%	12.7%	26.3%
Santa Ana	29.9%	9.3%	16.8%	49.0%	23.1%	23.1%
Piritu	19.9%	7.8%	9.4%	39.0%	12.1%	18.2%
Fernando de Peñalver	18.3%	6.3%	9.3%	33.5%	11.2%	15.6%
Freites	15.0%	5.3%	3.1%	28.9%	12.8%	14.9%
Independencia	5.0%	1.5%	2.8%	8.5%	5.3%	7.4%
Bolivar	7.6%	2.0%	4.5%	14.8%	5.1%	6.6%
Sotillo	4.0%	1.7%	1.6%	8.7%	2.6%	2.8%
Anaco	4.0%	2.3%	0.7%	9.5%	2.5%	2.5%
Guanta	0.7%	1.0%	0.0%	4.1%	0.9%	0.9%
Lic. Diego Bautista Urban	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Note: The results are sorted by 2020 estimates. *Oil-producing state.

Table 3.5: (Continued)

State and Municipality	Mean	Standard deviation	Minimum	Maximum	2014	2020
Monagas*	**25.7%**	**9.8%**	**9.8%**	**44.7%**	**18.8%**	**24.5%**
Uracoa	71.7%	8.0%	58.5%	85.8%	58.5%	69.3%
Aguasay	37.8%	18.8%	1.7%	67.8%	37.7%	51.5%
Acosta	47.1%	19.7%	1.1%	73.0%	34.2%	44.8%
Libertador	35.5%	8.5%	24.1%	53.7%	25.6%	35.9%
Caripe	28.5%	9.2%	10.0%	47.0%	23.4%	26.1%
Piar	19.6%	15.1%	0.0%	52.5%	6.7%	20.2%
Bolivar	26.1%	9.5%	11.6%	46.2%	18.3%	19.7%
Maturin	16.4%	4.7%	9.3%	27.5%	13.3%	16.8%
Sotillo	11.0%	2.1%	7.2%	15.5%	9.0%	13.8%
Punceres	22.0%	12.4%	3.9%	48.1%	13.1%	11.1%
Cedeño	14.0%	12.0%	0.1%	42.8%	4.4%	9.0%
Santa Barbara	4.1%	6.4%	0.0%	19.3%	0.0%	0.6%
Ezequiel Zamora	0.3%	0.7%	0.0%	2.5%	0.0%	0.0%
Yaracuy	**29.4%**	**7.2%**	**20.4%**	**47.0%**	**25.4%**	**24.5%**
Veroes	63.3%	11.6%	49.0%	96.1%	51.5%	51.1%
Bolivar	68.2%	12.6%	48.8%	86.7%	56.4%	49.1%
Manuel Monge	69.6%	20.4%	40.0%	97.0%	49.9%	43.3%
San Felipe	44.9%	9.4%	34.4%	76.7%	38.6%	39.2%
La Trinidad	46.3%	14.5%	33.3%	85.6%	41.6%	33.3%
Nirgua	33.1%	3.1%	29.0%	43.0%	33.8%	32.7%
Aristides Bastidas	10.8%	6.2%	3.0%	23.2%	15.8%	23.2%
Peña	19.5%	6.1%	12.2%	37.5%	12.2%	16.6%
Bruzual	14.0%	4.8%	8.6%	32.4%	12.7%	12.1%
Jose Antonio Paez	12.1%	3.7%	8.5%	26.2%	11.2%	10.7%
Urachiche	9.4 %	1.8%	6.8%	14.5%	11.2%	10.2%
Cocorote	6.0%	1.8%	3.1 %	9.8%	8.4%	8.4%
Independencia	9.4%	2.9%	5.7%	18.9%	7.4%	7.5%
Sucre	4.7%	1.5%	2.8%	10.1%	4.7%	5.6%
Aragua	**17.5%**	**4.1%**	**11.2%**	**26.5%**	**14.5%**	**18.3%**
Urdaneta	67.2%	5.3%	58.2%	76.9%	58.2%	65.1%
Camatagua	45.2%	8.0%	29.2%	64.2%	35.8%	45.1%
San Casimiro	35.9%	5.3%	28.5%	50.5%	31.0%	37.4%
San Sebastian	34.0%	4.8%	22.8%	47.7%	32.8%	33.7%
Tovar	41.2%	14.8%	18.9%	73.7%	22.9%	31.8%
Ocumare de la Costa de Oro	37.1%	11.5%	20.1%	60.8%	28.1%	30.5%
Santos Michelena	14.5%	6.9%	5.8%	26.8%	12.9%	26.4%
Santiago Mariño	14.8%	2.9%	9.2%	19.2%	13.9%	18.6%
Jose Rafael Revenga	3.8%	3.8%	0.2%	11.4%	3.7%	11.2%
Jose Felix Ribas	6.1%	2.9%	2.1%	14.5%	4.5%	7.6%
Mario Briceño Iragorry	2.7%	3.2%	0.0%	8.9%	5.2%	7.4%
Zamora	5.7%	1.3%	3.9%	9.7%	5.9%	6.7%
Girardot	3.9%	0.6%	2.2%	4.7%	4.2%	4.0%
Sucre	2.1%	1.5%	0.0%	6.7%	2.1%	2.3%
Bolivar	0.5%	0.6%	0.0%	1.9%	0.4%	1.8%
Francisco Linares	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Jose Angel Lamas	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Libertador	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Nueva Esparta	**13.8%**	**2.1%**	**11.6%**	**18.7%**	**12.2%**	**11.9%**
Tubores	100	0.0%	100.0%	100.0%	100.0%	100.0%
Peninsula de Macanao	17.8%	2.4%	14.2%	25.2%	16.5%	15.7%
Isla de Coche	32.6%	18.9%	13.3%	76.0%	17.1%	14.2%
Diaz	0.3%	0.2%	0.0%	0.7%	0.5%	0.6%
Gomez	0.1%	0.2%	0.0%	0.6%	0.0%	0.4%
Antolin del Campo	0.0%	0.1%	0.0%	0.3%	0.0%	0.0%
Arismendi	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Garcia	0.6%	0.6%	0.0%	2.1%	0.4%	0.0%
Maneiro	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Marcano	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%
Mariño	0.2%	0.2%	0.0%	0.5%	0.0%	0.0%
Miranda	**9.9%**	**2.3%**	**6.4%**	**15.0%**	**8.0%**	**10.0%**
Pedro Gual	45.5%	6.1%	36.1%	57.7%	36.9%	39.7%
Acevedo	40.3%	6.9%	30.6%	58.6%	30.6%	37.2%
Paez	31.3%	5.2%	23.2%	43.5%	23.2%	26.8%
Paz Castillo	11.5%	5.1%	4.2%	22.4%	11.0%	17.3%
Andres Bello	11.9%	4.8%	1.5%	22.6%	6.4%	13.9%
Brion	13.6%	1.9%	9.8%	18.7%	11.1%	13.8%
Buroz	11.9%	4.5%	5.9%	22.1%	7.2%	10.7%
Lander	7.8%	1.1%	5.2%	10.0%	7.3%	8.6%
Zamora	9.6%	2.5%	6.0%	16.3%	8.3%	8.5%
Independencia	8.1%	0.8%	6.8%	9.8%	6.9%	7.7%
Guaicaipuro	3.5%	1.4%	1.4%	6.0%	3.9%	6.0%
Urdaneta	3.6%	1.6%	0.4%	6.8%	3.2%	5.0%
Plaza	2.5%	1.4%	0.7%	5.3%	3.2%	4.7%
Simon Bolivar	4.1%	1.1%	2.2%	6.4%	4.0%	4.3%
Sucre	1.5%	1.1%	0.0%	3.4%	2.9%	3.2%
Cristobal Rojas	1.2%	1.4%	0.0%	3.4%	1.8%	2.7%
Chacao	0.2%	0.3%	0.0%	0.8%	0.7%	0.5%
El Hatillo	0.3%	0.3%	0.0%	1.2%	0.4%	0.3%
Baruta	0.1%	0.3%	0.0%	1.0%	0.0%	0.0%
Carrizal	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Los Salias	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Note: The results are sorted by 2020 estimates. *Oil-producing state.

Table 3.5: (Continued)

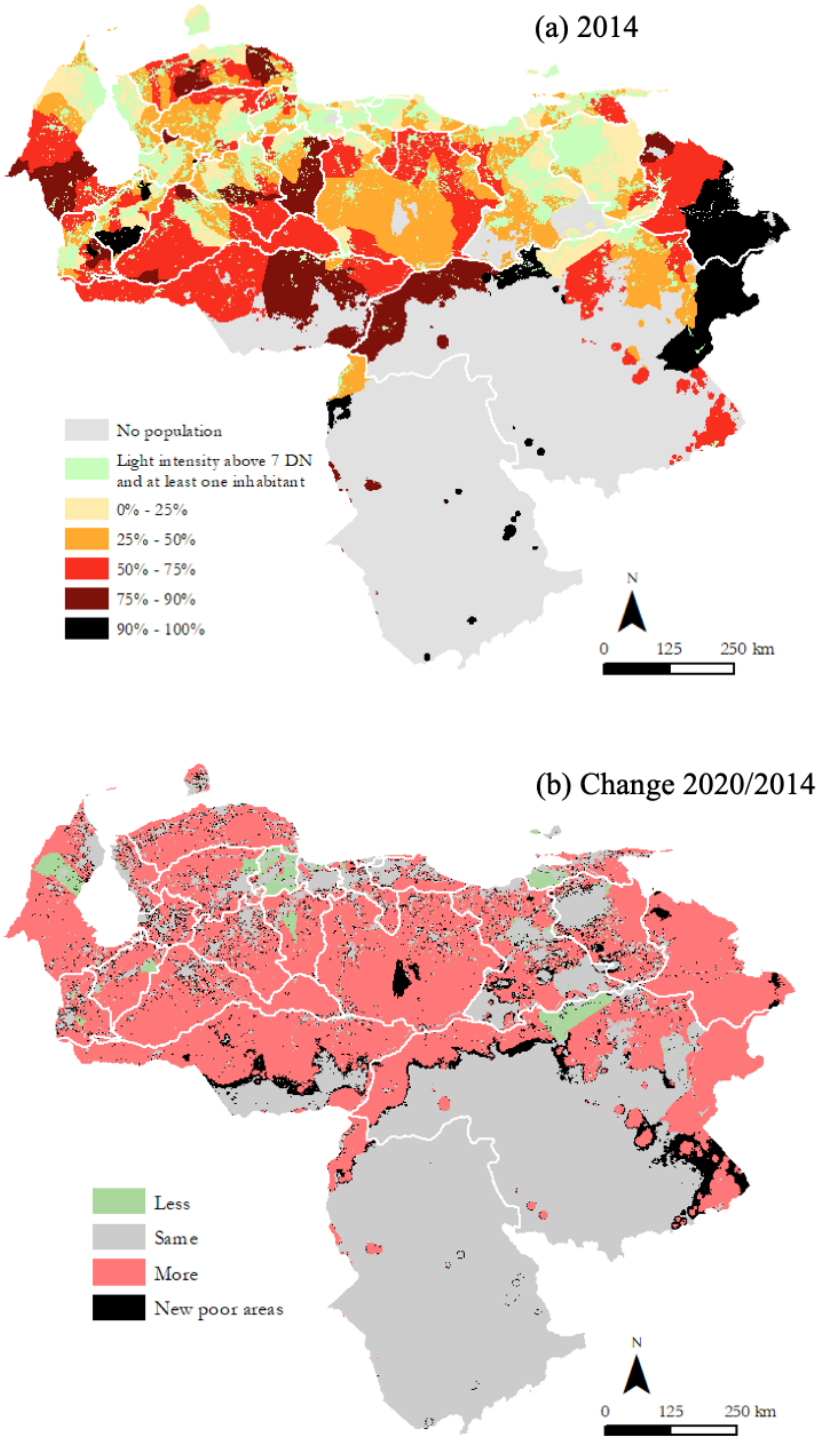
State and Municipality	Mean	Standard deviation	Minimum	Maximum	2014	2020
Vargas	**9.3%**	**2.4%**	**6.6%**	**14.7%**	**7.0%**	**7.6%**
Vargas	9.3%	2.4%	6.6%	14.7%	7.0%	7.6%
Carabobo	**5.2%**	**2.2%**	**2.0%**	**9.9%**	**5.8%**	**7.0%**
Bejuma	23.4%	7.0%	14.6%	37.1%	29.4%	31.8%
Montalban	14.4%	7.2%	1.0%	28.9%	17.3%	19.9%
Carlos Arevalo	10.6%	3.3%	6.1%	18.7%	7.3%	12.2%
Puerto Cabello	7.3%	3.4%	1.9%	16.0%	8.4%	8.0%
Juan Jose Mora	6.9%	1.9%	3.6%	12.6%	7.2%	7.2%
San Joaquin	2.2%	2.2%	0.0%	6.7%	3.9%	6.7%
Miranda	5.6%	1.9%	1.4%	9.2%	4.7%	5.0%
Diego Ibarra	1.0%	1.4%	0.0%	4.1%	1.7%	3.9%
Naguanagua	1.0%	1.3%	0.0%	3.6%	1.9%	3.5%
Guacara	1.7%	1.0%	0.1%	3.4%	2.5%	3.0%
Valencia	2.1%	1.0%	0.7%	4.7%	1.0%	1.7%
Libertador	0.7%	0.7%	0.0%	1.7%	0.7%	1.4%
San Diego	0.4%	0.5%	0.0%	1.2%	0.9%	1.2%
Lago Valencia	0.0%	0.1%	0.0%	0.3%	0.0%	0.0%
Los Guayos	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Distrito Capital	**0.1%**	**0.1%**	**0.0%**	**0.2%**	**0.2%**	**0.1%**
Libertador	0.1%	0.1%	0.0%	0.2%	0.2%	0.1%

Figure 3.8 shows how the rural poverty is spatially distributed in Venezuela for 2014 (panel a), as well as the change of rural poverty rates from 2014 to 2020 (panel b). In 2014, the higher rural poverty rates locate in the south and central Merida, central Tachira, and the Guayana region (Amazonas, Bolivar, and Delta Amacuro). Apure, Barinas, Falcon, and Zulia also show moderate to high rates in most of their territory. The Capital region (Miranda, Vargas, and the Distrito Capital), the Eastern region (Anzoategui, Monagas, and Sucre), Aragua, Carabobo, Lara, Yaracuy, and the central north of Portuguesa have moderate to low rates. The Colombia-Venezuela border to the west of Tachira shows light intensity above the threshold, making difficult the identification of rural poverty; however, I cannot rule out the presence of urban poverty.

Rural poverty is increasing during the collapse, and new rural poor areas are appearing between 2014 and 2020 in clusters, mostly surrounding municipalities with moderate to high poverty rates. However, rural poverty does not follow similar paths across all states and municipalities. For example, except for Nueva Esparta, the southwest of Sucre, Yaracuy, and some municipalities in Bolivar, Cojedes, Lara, Merida, Tachira, and Zulia, the rest of the country's territory experienced an intensification of rural poverty.[9] These heterogeneities may be caused by the electricity rationing frequently imposed since 2008, the regime policy decisions prioritizing the power demand of main urban cores, and the mismanagement of the electricity sector.

[9]For details, see Table 3.5.

Figure 3.8: Mapping the Rural Poverty Rate in Venezuela



Source: Own calculations.

3.5 Conclusion

The lack of local poverty data is a significant limitation to setting development policies in emerging countries, and this chapter tries to close that gap for rural Venezuela. Here I use an alternative approach based on remote sensing techniques to generate subnational rural poverty data for Venezuela, a country facing political turmoil and severe socio-economic crisis. In this case, I use DMSP and VIIRS nighttime light imagery and two spatially distributed population datasets, ORNL and WorldPop, to estimate rural poverty rates at the state and municipality levels from 2000 to 2020.

The harmonization procedure between DMSP and VIIRS products identified a light intensity threshold for low values associated with poor rural areas in Venezuela ($DN = 7$). This threshold allows for obtaining the poverty estimates. The chapter also verifies how the WorldPop data led to more precise and consistent results over time, suggesting its use for further geospatial analysis for Venezuela. With this technical approach, the findings identify Amazonas, Apure, and Delta Amacuro as the top three states with the highest rural poverty rate in 2020, and Carabobo, Vargas, and the Distrito Capital as the bottom three.

This chapter confirms subnational heterogeneities in the results but with broad common trends. The satellite sensitivity reveals volatility during 2003–2004 and, in most cases, an absolute peak in 2003 (a year of significant political instability). There was a switching period in 2013–2014 (beginning the current Venezuelan recession), from which rural poverty intensified and new rural poor areas appeared across the country, surrounding municipalities with moderate to high poverty rates. Except for Nueva Esparta, the southwest of Sucre, Yaracuy, and some municipalities in Bolivar, Cojedes, Lara, Merida, Tachira, and Zulia, the rest of the country experienced an intensification of rural poverty. These results suggest poverty clustering and that more Venezuelans have sunk into darkness in recent years.

Since 2008 the Venezuelan regime is frequently imposing electricity rationing regionally to prioritize the energy demand of main urban cores (such as the Distrito Capital), which may contribute to these subnational heterogeneities. On the other hand, mismanagement of the electricity sector prevails while a narrow political criterion still drives the planning and execution of official generation capacity projects. Unfortunately,

this context highlights the lack of effective territorial intelligence and the need for a development agenda that is politically unbiased and based on reliable, independent subnational data.

The energy crisis is underway and even escalated with the 2019 national blackout, which had humanitarian costs across the healthcare, water supply and public transport systems, and significant disrupting effects on the commercial and retail sectors and oil production (Sabatini and Patterson 2021). This context yielded more electricity rationing in Western states (Apure, Barinas, Merida, Tachira, Trujillo, and Zulia), contemplating official energy interruption of four hours in each state, likely impacting the latest poverty results.[10] Now, inhabitants of those states reported sustained economic degradation and actual energy cuts of about 6–12 hours or even days (Prensa Aula Abierta 2020; Venezuelan Observatory of Public Services 2021).

The Venezuelan context warns about the current need to reform its electricity sector. A significant step toward it is to build up local data to support eventual recovery efforts and poverty alleviation programs and allow impact evaluations. That is the main contribution of this chapter, which findings also encourage future research zooming in on regional inequalities and identifying national binding constraints in infrastructure development to promote economic growth.

[10]The National Electricity Corporation of Venezuela explained that these "load management plans" are caused by low levels in the Uribante reservoir.

Chapter 4

Measuring Regional Inequality in the Andean Countries: A Multiple-Stage Nested Theil Decomposition Using Night Light Emissions

4.1 Introduction

Since the early 2000s and until 2019, more than 45 million people have risen out of income poverty across Latin America and the Caribbean (LAC), of which about 15 million people lived in the Andean Region (Bolivia, Colombia, Ecuador, Peru, and Venezuela).[1] However, the COVID-19 pandemic widely dominated 2020 and caused severe impacts on socioeconomic indicators, pulling back growth forecasts and significantly increasing poverty-vulnerable populations and income inequality (Castilleja 2020; De la Cruz et al. 2020; Lanjouw and Tarp 2021, Andrian and Manzano 2023). As a result, about 48.3 million people in the Andean region suffered income poverty by 2021 (a fifth being extremely poor), 32 percent and 8 percent more than in 2012 and 2019, respectively.[2]

Poverty is closely related to regional inequality.[3] Economic and social opportunities are often unequally distributed across regions, and poverty tends to concentrate in areas with lower economic development and resource access. In many cases, these regions are rural areas (Rodríguez-Pose and Hardy 2015; Economic Commission for Latin America and the Caribbean, ECLAC, 2018; Kharas et al. 2020; Maldonado 2023), not densely populated and where agriculture is the primary source of income, and urban slums, weak and densely packed housing units where inhabitants have limited access to essential services (such as clean water, sanitation, education, and healthcare) and fewer job opportunities (Stampini et al. 2015). For example, about 38 percent of the rural population in the Andean region faced income poverty in 2021 (11.3 million people), while 31 percent were urban poor (37 million people). Regional inequality can also lead to migration as people move from poorer regions to more affluent ones in search of better opportunities. This situation may lead to structural rather than contextual issues, a local brain drain, loss of economic activity, and uneven regional growth, further exacerbating poverty in the poorer regions and driving intergenerational transmission of

[1] For details, see the Economic Commission for Latin America and the Caribbean's Statistical Database (2022, January): https://statistics.cepal.org/portal/cepalstat/dashboard.html?lang=en. Due to data availability constraints for Venezuela, its poverty figures are based on the National Survey of Living Conditions for Venezuela and from Maldonado (2023).

[2] Venezuela influences these figures up, although the trend remains. Excluding Venezuela, 33.3 million people were income poor by 2021 (about 10 million of which were extremely poor), 10 percent more than in 2012 and 15 percent more than in 2019.

[3] Regional inequality refers to economic, social, or political disparities between different regions within a country or geographic area.

poverty (Bird 2013).

Governments face a first challenge when addressing inequality in their countries: what is the most suitable geographic level to focus their policy efforts to promote equality objectives? This is a fundamental and critical aspect of policy design and implementation. Choosing an appropriate local scale of action may significantly influence the precision and effectiveness of initiatives aimed at tackling inequality. Policymakers can tailor interventions to address the specific dynamics and challenges in a particular region and promote inclusivity and participatory governance while considering the preferences and needs of specific geographical levels.

In this sense, this chapter uses satellite-recorded nighttime lights and gridded population datasets to shed light on the appropriate geographical level for implementing initiatives to tackle inequality in each country within the Andean region. I measure inequality accounting for their lowest (smallest)-level administrative divisions and examine how it may have varied from 2012 to 2021. In the best scenario, the analytical approach involves decomposing inequality following a three-stage nested Theil index decomposition method using night light emissions as a proxy of economic wealth.[4] This study confirms a decrease in overall wealth inequality for the Andean region throughout the period (primarily driven by a decline in between-country inequality) and an increase in the relative importance of within-country inequality. It also identifies spatial heterogeneities by country, particularly in Venezuela, the only one in the Andean region where inequality increased.

Nighttime lights from satellite imagery are often used in regional analyses to track economic activity and economic development (Dai et al. 2017; Wang et al. 2019; Andrade-Núñez and Aide 2020; Gibson and Boe-Gibson 2021; Maldonado 2022; McCord and Rodriguez-Heredia 2022), and socioeconomic and political outcomes (Hodler and Raschky 2014; Bruederle and Hodler 2018; Ferreira 2018; Jagnani and Khanna 2020; Maldonado 2023). These studies rely on the assumption that night light emissions implicitly capture relevant information about spatial heterogeneity and human impact on a local level.[5]

[4]Econometrician Henri Theil introduced the Theil decomposition method in the 1960s, based on the Theil index, an entropy measure of inequality. The decomposition considers both the inequality between groups and the inequality within groups. It involves partitioning the total variation in the dependent variable into different components, each representing a different factor's contribution to the overall variation.

[5]Refer to Chapter 2 and Chapter 3.

Regional inequality and light emissions are deeply interconnected. The availability and consumption of energy, such as electricity, is fundamental to modern living and a key input to industrial and trade activity, while its absence can be a significant constraint on income and development (Moss et al. 2020). Lights can also indicate inequality. Wealthier areas have more lighting at night, while poorer and more deprived areas have less. Moreover, urban areas concentrate electricity and energy infrastructure emitting an intensity of light likely captured by sensors aboard satellites. In contrast, electrification is inferior in rural areas or agricultural regions, mainly because it is far from national grids, thus leading to less intense lighting (Keola et al. 2015; Ferreira 2018; Smith and Wills 2018; Maldonado 2023). In this sense, uneven light emissions across the territory may signaling different degrees of development and economic activities.

Conventional methods of measuring inequality typically incorporate data from national accounts, administrative records, household surveys, or a mix. However, these sources are susceptible to discrepancies in design and inconsistent accessibility (Dahl et al. 2011; Burkhauser et al. 2012; Deaton 2016; Carr and Wiemers 2018; Galimberti et al. 2023). For example, administrative data can have missing values or be outdated (Courtemanche et al. 2019), national accounts are not designed to generate poverty estimates, and surveys tend to under-sampling richer households likely due to forgetfulness, temporal misplacement, or misclassification (Deaton 2005; Lynn et al. 2012). Ayala et al. (2022) also confirm significant differences in the level and structure of inequality across administrative and survey data, particularly in the tails of the income distribution. Tax avoidance and evasion are also issues in administrative data. Tax records do not account for informal sources and may be limited by fiscal manipulation strategies and income reporting rules (Alstadsæter et al. 2019; Galbraith 2019; Meyer and Mittag 2021).

Analyzing regional inequality based on alternative sources and methods may generate new outputs or complement existing results. A feasible but insufficiently exploited approach is to estimate light-based inequality measures. Geospatial source data have consistent coverage worldwide and are less prone to transitory income shocks. The intensity of light emissions can also reveal rural-urban territorial features (Zhao et al. 2019; Li et al. 2020), capture specificities beyond reported income (such as informal activities, access to electricity, and infrastructure), and act as independent novel data with predictive power to audit and reduce measurement errors from surveys and national

accounts (Henderson et al. 2012; Nordhaus and Chen 2012; Chen 2016; Pinkovskiy and Sala-i-Martin 2016; Maldonado 2022; Martinez 2022).

Satellite imagery is slowly gaining ground as a proxy of income and wealth for estimating regional inequality in developed and developing countries. Mveyange (2015) found a significant positive relationship between regional inequality calculated through nighttime lights and income in Africa. Lessmann and Seidel (2017) used luminosity data to estimate regional income inequality across 180 countries at the first subnational administrative level from 1992 to 2012. Their study revealed that predicted income is a more precise representation of real figures than nominal figures, which is crucial in assessing regional disparities. Weidmann and Schutte (2017) also demonstrated the accuracy of light data in predicting local economic wealth for 39 of the world's least-developed countries.

More recently, Galimberti et al. (2023), Weidmann and Theunissen (2021), and Andreano et al. (2021) have contributed to that growing literature. Galimberti et al. (2023) confirmed the superiority of using satellite data on nighttime lights and spatially distributed population data to measure economic inequality within developing countries. Weidmann and Theunissen (2021) focused on African countries and suggested that nighttime light can be an alternative to survey-derived data when estimating local disparities. Furthermore, Andreano et al. (2021) analyzed a panel of 20 countries in Latin America and the Caribbean at the first subnational administrative level between 2000 and 2013. Their findings suggest that nighttime lights could be a critical source of information for deriving spatially disaggregated and continuous-time calculations of inequality indices.

The contribution of this chapter is threefold. First, there is a lack of studies using remote sensing data from satellite imagery to expose regional inequality accounting for multiple subnational levels in each Andean country. In fact, this study is the first one to use satellite-recorded light data to study inequality, focusing solely on the Andean region. There are studies for each Andean country using light data but analyzing map urbanization dynamics and estimating economic or social indicators. Parés-Ramos et al. (2013) used nighttime lights to analyze spatial patterns of urban development between 1992 and 2009 in the major cities of Bolivia, Colombia, Ecuador, and Peru. I also came across a blog post by Andersen et al. (2023, January) presenting results of night light

intensity by municipalities of Bolivia. In Ecuador, Cabrera-Barona et al. (2020) mapped urban representations of the Ecuadorian Amazon using a geospatial approach, and Mejía Juárez (2020) analyzed the evolution of urban uses using the magnitude and intensity of nighttime light data. In Peru, Seminario and Palomino (2022) estimated subnational GDP from 1993 to 2018. Zhang et al. (2020) carried out a spatiotemporal analysis relating light emissions with multiple variables to evaluate the socioeconomic crisis that suffers Venezuela, while Maldonado (2023) actually combined the absence of high intensity of lights and gridded population datasets to estimate 2000-2020 rural poverty rates at different geographic levels of Venezuela. Furthermore, using satellite data as an alternative source is particularly important for analyzing developing countries, which are often limited or lack robust subnational indicators to diagnose regional disparities.

Second, this chapter uses the latest consistently processed annual series of nighttime lights -version 2.1- collected by the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument at its highest available spatial resolution (approximately 500 meters at the equator) from 2012 to 2021. Earlier studies on local inequality using satellite data often rely on yearly light imagery from the U.S. Air Force Defense Meteorological Satellite Program (DMSP)-Operational Linescan System (publicly available from 1992 to 2013); however, VIIRS data are superior to DMSP for mapping nighttime lights (Elvidge et al. 2013). VIIRS spatial resolution provides more detailed information than DMSP. In addition, unlike VIIRS data, DMSP sensors lack linearity, suffer from frequent saturation, and have no onboard calibration.[6] Therefore, results using VIIRS data ensure a gain in reliability and precision on subnational outcomes, keeping us (for a while) at the forefront of other studies on regional inequality.

Third, a comprehensive analysis of inequality measures requires considering both spatial and temporal dimensions (Khan and Siddique 2021). This chapter accounts for both dimensions. Here, I analyze spatiotemporal heterogeneities at the lowest-level administrative divisions of each Andean country. This chapter follows a multiple-stage nested Theil decomposition method covering ten years, 2012-2021. This approach allowed us to find noteworthy particularities. For example, the chapter verifies a decline in wealth inequality over the past decade in Bolivia, Colombia, and Peru. It finds that Bolivia's

[6]This means that assigned values of intensity of DMSP light data are not comparable from one year to another.

reduction in inequality can be attributed to the decline in disparities between provinces. Similarly, Colombia experienced a decrease in inequality within municipalities, and Peru witnessed a decline within districts. In contrast, the inequality components in Ecuador and Venezuela exhibit a more balanced contribution to overall inequality. And, while Ecuador does not show a significant change in overall inequality during the period, the inequality increase in Venezuela is primarily driven by changes in the disparity between all geographic subgroups.

As far as we know, recent studies have yet to apply multiple-stage nested geographic disaggregations for Andean countries, although some exercises exist for Chile and China. A decade ago, Paredes et al. (2012) empirically proposed a Theil decomposition for Chile at the regional, provincial, and county levels, assuming household incomes as the lowest level observations. In addition,Wu et al. (2018) analyzed inequality in research funding in China at the individual researcher level by university-institute subgroup.

The recent global policy agenda acknowledges the importance of including different levels of geography to understand multiple degrees of development within and among countries (Sachs et al. 2022). Reducing inequality is in ten out of 17 goals adopted by all United Nations (UN) member States in 2015 as part of the 2030 Agenda for Sustainable Development. Floerkemeier et al. (2021) further emphasize that inequality measures are sensitive to the choice of geographical scale, and spatially targeted policies could help foster more inclusive regional development outcomes and aggregate sustainable economic growth. This study is a step towards promoting a spatial perspective into policy design and implementation in Andean countries, an approach that could play a vital role in achieving equitable development goals at multiple geographic levels.

This chapter could be beneficial for two main strands of inequality studies. First, It complements literature using geospatial data to measure national or subnational inequality (Lessmann and Seidel 2017; Gilliland et al. 2019; Haithcoat et al. 2021; Mirza et al. 2021; Rabiei-Dastjerdi and Matthews 2021; Puttanapong et al. 2022; Galimberti et al. 2023). Second, It also expands existing studies on inequality based on decomposition methods (Morduch and Sicular 2002; Akita 2003; Shorrocks and Wan 2005; Elbers et al. 2008; Paredes et al. 2012; Wu et al. 2018; Sinha et al. 2022).

The chapter is structured as follows. Section 4.2 summarizes the analytical framework describing the conventional one-stage Theil decomposition method and its extension to

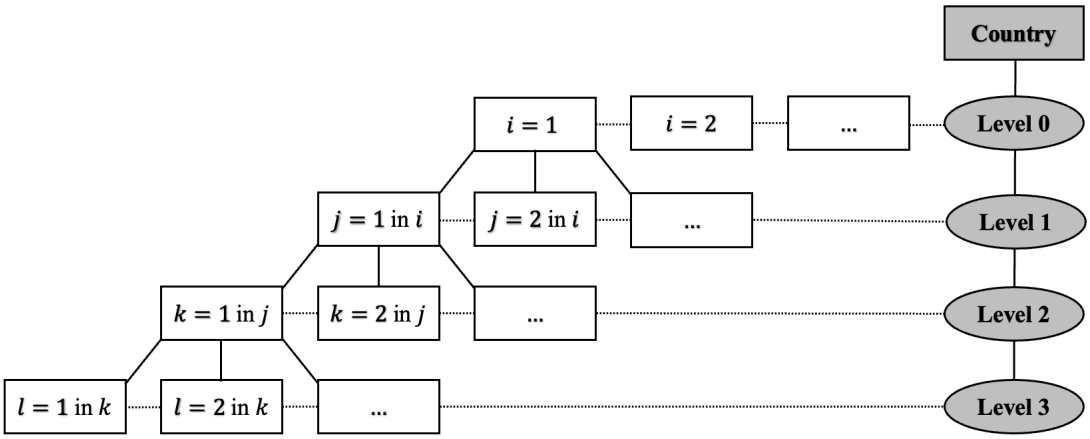
the two-stage nested and three-stage nested Theil decomposition method. Section 4.3 describes the study area and data sources. Section 4.4 presents stylized facts of night light data for the sample and main results. Section 4.5 concludes.

4.2 Methodology

The Theil index[7] was designed to measure income and wealth disparity within a country. As indicated by Akita (2003), the index comprises two variants: the Theil T index, which employs income proportions as weights and tends to be more responsive to changes in more affluent areas, and the Theil L index, which uses population proportions and thus is particularly responsive to changes among poorer areas. Here, the framework is based on the Theil T index.

Figure 4.1 shows four layers of geographic levels and sublevels nested in a country, where the last layer (*Level* 3) contains the underlying unit l to measure regional income inequality.

Figure 4.1: Four-Level Hierarchical Structure



The country-level Theil T index (overall inequality) will depend on the number of levels defined a priori for the hierarchical structure, such as T_1 is the Theil index assuming a one-level hierarchical structure, T_2 assuming a two-level hierarchical structure, T_3 for

[7] Developed by Theil (1967) using principles of information theory.

a three-level hierarchical structure, or T_4 for a four-level hierarchical structure. These are given by the following equations:

$$T_1 = \sum_i \left(\frac{Y_i}{Y}\right) Log\left(\frac{Y_i/Y}{N_i/N}\right) \tag{4.1}$$

$$T_2 = \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) Log\left(\frac{Y_{ij}/Y}{N_{ij}/N}\right) \tag{4.2}$$

$$T_3 = \sum_i \sum_j \sum_k \left(\frac{Y_{ijk}}{Y}\right) Log\left(\frac{Y_{ijk}/Y}{N_{ijk}/N}\right) \tag{4.3}$$

$$T_4 = \sum_i \sum_j \sum_k \sum_l \left(\frac{Y_{ijkl}}{Y}\right) Log\left(\frac{Y_{ijkl}/Y}{N_{ijkl}/N}\right) \tag{4.4}$$

where Y is the total income at the country-level, Y_i is the total income of i (in *Level* 0), Y_{ij} is the total income of j (in *Level* 1) in i, Y_{ijk} is the total income of k (in *Level* 2) in j in i, Y_{ijkl} is the total income of l (in *Level* 3) in k in j in i, N is the total population at the country-level, N_i is the total population of i (in *Level* 0), N_{ij} is the total population of j (in *Level* 1) in i, N_{ijk} is the total population of k (in *Level* 2) in j in i, and N_{ijkl} is the total population of l (in *Level* 3) in k in j in i. Therefore, $Y = \sum_i Y_i = \sum_i \sum_j Y_{ij} = \sum_i \sum_j \sum_k Y_{ijk} = \sum_i \sum_j \sum_k \sum_l Y_{ijkl}$ and $N = \sum_i N_i = \sum_i \sum_j N_{ij} = \sum_i \sum_j \sum_k N_{ijk} = \sum_i \sum_j \sum_k \sum_l N_{ijkl}$.

The Theil index is additively decomposable, meaning that the overall national inequality can be expressed as the sum of within-groups and between-groups inequality components.[8] It quantifies the degree to which the structure in the distribution of income across groups deviates from the distribution of population across those same groups. When the structures are the same, there is an equal income distribution for all underlying units; thus, each group has the same share of income as its population share, and the Theil index equals zero (minimum value). Similarly, if a specific group

[8]Theil indices satisfy convenient properties as a measure of regional income inequality: mean independence (the index remains unchanged if every group's income changes by the same proportion), population-size independence (the index remains unchanged if the population in each group change by the same proportion), and the Pigou-Dalton principle of transfers (any income transfer from a richer to a poorer group that does not reverse their relative ranks in income reduces the value of the index). For details, see Bourguignon (1979), Shorrocks (1980), and Akita (2003).

has the same share of income and population, its relative contribution to the index is zero. On the other hand, the larger the index, the larger the inequality. Groups with higher shares of income relative to the population contribute positively to the Theil index, and those with lower shares of income than the population contribute negatively. By definition, the positive contributions will always be higher than the negative ones, so the index or its components will always be positive overall or zero.[9]

Let's consider a simplified case of a two-level hierarchical structure of a country, ignoring lower nested levels, where *Level* 1 encloses the underlying unit j. In this case, overall regional income inequality can be measured by Equation 4.2. If we define T_i to measure Theil T indices within i for *Level* 0 as

$$T_i = \sum_j \left(\frac{Y_{ij}}{Y_i}\right) Log\left(\frac{Y_{ij}/Y_i}{N_{ij}/N_i}\right) \tag{4.5}$$

then, the within-0 inequality component (T_{W0}) is a weighted average of Equation 4.5 using income shares as weights at the *Level* 0 as in Equation 4.6,

$$T_{W0} = \sum_i \left(\frac{Y_i}{Y}\right) T_i = \sum_i \left(\frac{Y_i}{Y}\right) \sum_j \left(\frac{Y_{ij}}{Y_i}\right) Log\left(\frac{Y_{ij}/Y_i}{N_{ij}/N_i}\right) \tag{4.6}$$

Noting that Equation 4.7 measures the between-0 inequality component (T_{B0}) of the country or the inequality between subgroups in *Level* 0,

$$T_{B0} = \sum_i \left(\frac{Y_i}{Y}\right) Log\left(\frac{Y_i/Y}{N_i/N}\right) \tag{4.7}$$

then, the Equation 4.2 can be decomposed into,

$$\begin{aligned} T_2 &= \sum_i \left(\frac{Y_i}{Y}\right) \sum_j \left(\frac{Y_{ij}}{Y_i}\right) Log\left(\frac{Y_{ij}/Y_i}{N_{ij}/N_i}\right) + \left(\frac{Y_i}{Y}\right) \sum_i \left(\frac{Y_i}{Y}\right) Log\left(\frac{Y_i/Y}{N_i/N}\right) \\ &= \sum_i \left(\frac{Y_i}{Y}\right) T_i + \sum_i \left(\frac{Y_i}{Y}\right) Log\left(\frac{Y_i/Y}{N_i/N}\right) = \sum_i \left(\frac{Y_i}{Y}\right) T_i + T_{B0} = T_{W0} + T_{B0} \end{aligned} \tag{4.8}$$

Equation 4.8 represents the conventional one-stage Theil decomposition.

[9] For guidance about intuitive interpretations and analytical applications of the Theil index, see Conceição and Ferreira (2000).

I can also assume a three-level hierarchical structure, adopting k as the underlying unit in *Level* 2, to get a two-stage nested Theil decomposition.[10] In this case, overall regional income inequality can be measured by Equation 4.3.

By analogy, we can further decompose the Theil T indices of Equation 4.5 into,

$$T_i = \sum_j \left(\frac{Y_{ij}}{Y_i}\right) T_{ij} + \sum_j \left(\frac{Y_{ij}}{Y_i}\right) Log\left(\frac{Y_{ij}/Y_i}{N_{ij}/Ni}\right) = \sum_j \left(\frac{Y_{ij}}{Y_i}\right) T_{ij} + T_{Bi} \tag{4.9}$$

where T_{Bi} represents the inequality between each j in i. And, if I define T_{ij} to measure Theil T indices within j in i for *Level* 1,

$$T_{ij} = \sum_k \left(\frac{Y_{ijk}}{Y_{ij}}\right) Log\left(\frac{Y_{ijk}/Y_{ij}}{N_{ijk}/N_{ij}}\right) \tag{4.10}$$

By substituting Equation 4.9 in Equation 4.8, and using Equation 4.10, I obtain Equation 4.11.

$$\begin{aligned} T_3 &= \sum_i \left(\frac{Y_i}{Y}\right) \left[\sum_j \left(\frac{Y_{ij}}{Y_i}\right) T_{ij} + T_{Bi}\right] + T_{B0} \\ &= \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) T_{ij} + \sum_i \left(\frac{Y_i}{Y}\right) T_{Bi} + T_{B0} \\ = \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) T_{ij} &+ \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) Log\left(\frac{Y_{ij}/Y_i}{N_{ij}/N_i}\right) + T_{B0} = T_{W1} + T_{B1} + T_{B0} \end{aligned} \tag{4.11}$$

where T_{B1} represents the inequality between subgroups in *Level* 1 (between-1 inequality component), and T_{W1} is a weighted average of the within-group j Theil indices T_{ij} (within-1 inequality component).

Furthermore, I can now assume the four-level hierarchical structure, where *Level* 3 comprises the underlying unit l. In this case, Equation 4.4 measures the overall regional income inequality, and I should now decompose the Theil T indices of Equation 4.10,

[10] For details, see Akita (2003).

obtaining thus Equation 4.12.

$$T_{ij} = \sum_k \left(\frac{Y_{ijk}}{Y_{ij}}\right) T_{ijk} + \sum_k \left(\frac{Y_{ijk}}{Y_{ij}}\right) Log\left(\frac{Y_{ijk}/Y_{ij}}{N_{ijk}/N_{ij}}\right) = \sum_k \left(\frac{Y_{ijk}}{Y_{ij}}\right) T_{ijk} + T_{Bij} \quad (4.12)$$

where T_{Bij} is the inequality between each k in j in i.

If I define T_{ijk} to measure Theil T indices within k in j in i for *Level* 2,

$$T_{ijk} = \sum_l \left(\frac{Y_{ijkl}}{Y_{ijk}}\right) Log\left(\frac{Y_{ijkl}/Y_{ijk}}{N_{ijkl}/N_{ijk}}\right) \quad (4.13)$$

Finally, by substituting Equation 4.12 in Equation 4.11, and using Equation 4.13, I have

$$\begin{aligned} T_4 &= \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) \left[\sum_k \left(\frac{Y_{ijk}}{Y_{ij}}\right) T_{ijk} + T_{Bij}\right] + \sum_i \left(\frac{Y_i}{Y}\right) T_{Bi} + T_{B0} \\ &= \sum_i \sum_j \sum_k \left(\frac{Y_{ijk}}{Y}\right) T_{ijk} + \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) T_{Bij} + \sum_i \left(\frac{Y_i}{Y}\right) T_{Bi} + T_{B0} \\ &= \sum_i \sum_j \sum_k \left(\frac{Y_{ijk}}{Y}\right) T_{ijk} + \sum_i \sum_j \sum_k \left(\frac{Y_{ijk}}{Y}\right) Log\left(\frac{Y_{ijk}/Y_{ij}}{N_{ijk}/N_{ij}}\right) \\ &\quad + \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) Log\left(\frac{Y_{ij}/Y_i}{N_{ij}/N_i}\right) + T_{B0} = T_{W2} + T_{B2} + T_{B1} + T_{B0} \end{aligned} \quad (4.14)$$

where T_{B2} is the inequality between subgroups in *Level* 2 (between-2 inequality component), and T_{W2} is the weighted average of the within-group k Theil indices T_{ijk} (within-2 inequality component). Equation 4.14 represents the three-stage nested Theil decomposition.

4.3 Study Area and Data

This chapter carries out a multiple-stage nested Theil decomposition method to present results of wealth inequality for Bolivia, Colombia, Ecuador, Peru, and Venezuela between 2012 and 2021, using night light emissions as a proxy for wealth and their respective

lowest-level administrative divisions as the underlying regional unit.

The geographic levels considered by country are given in Table 4.1 and correspond to a four-level hierarchical structure, except for Colombia, which has a three-level structure.

Table 4.1: Geographic Levels by Andean Country

Country	*Level* 0	*Level* 1	*Level* 2	*Level* 3
Bolivia	Region 3	Department 9	Province 112	Municipality 339
Colombia	Region 5	Department 32+1 (Distrito Capital)	Municipality 1,122	
Ecuador	Region 4+1 (Non-delimited)	Province 24+1 (Non-delimited)	Canton 221+3 (Non-delimited)	Parish 1,040
Peru	Region 3	Department 24+1 (Constitutional Province of Callao)	195+1 (Constitutional Province of Callao)	District 1,873
Venezuela	Region 9	State 23+1 (Distrito Capital)	Municipality 335	Parish 1,134

Note: *Level* 0 corresponds to natural or physiographic regions (except political-administrative regions for Venezuela). The number of geographic levels may differ from official major administrative areas due to the availability of geospatial data.

I use shapefiles of subnational administrative boundaries from the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Common Operational Datasets to aggregate gridded data to divisions in *Level* 1, *Level* 2, and *Level* 3. In addition, I create regional divisions, *Level* 0, based on predefined physiographic features by country (or political-administrative regions for Venezuela). Adding *Level* 0 allows us to obtain more detailed inequality components while accounting for natural differences in land surface and climate, which lead to diversity, for example, in crops.

The Oak Ridge National Laboratory, ORNL (2023, February) provides the global population distribution data: LandScan Global Population database.[11] This is an ambient -average over 24 hours- gridded population data at 30 arc-second spatial resolution (approximately a cell area of 1x1 kilometers). Here, LandScan seems a suitable choice. Yin et al. (2021) conducted an accuracy assessment of four commonly used gridded population data products, including LandScan.[12] They concluded that LandScan performs best regarding spatial fineness and estimated errors. In this chapter, I aggregate LandScan adding it up throughout the different geographic levels by country.

[11]The files are freely available at the online repository of the ORNL: https://landscan.ornl.gov

[12]In particular, Yin et al. (2021) cross-compared global gridded population datasets such as the Gridded Population of the World (GPW), Global Human Settlement Population Grid (GHS-POP), WorldPop, and LandScan.

Regarding the light emissions data, I use the newly consistently processed time series -version 2.1- of annual global low-light imaging data captured by the VIIRS Day/Night Band (DNB) onboard the Suomi National Polar-Orbiting Partnership satellite platform. The intensity of light is measured by the DNB radiance at night with unit: nano watts per square centimeter per steradian, $nW/(cm^2/sr)$. The Earth Observation Group of the Payne Institute for Public Policy at the Colorado School of Mines processes these light raster products, currently available from 2012 to 2021 (Earth Observation Group 2022, June).[13] Looking for pre-filtered spatially refined data, I use VIIRS annual masked average radiance products (Elvidge et al. 2021) at its highest spatial resolution of 15 arc-seconds or approximately 500 meters at the equator, which is particularly important when aggregating to a detailed-geographic level.

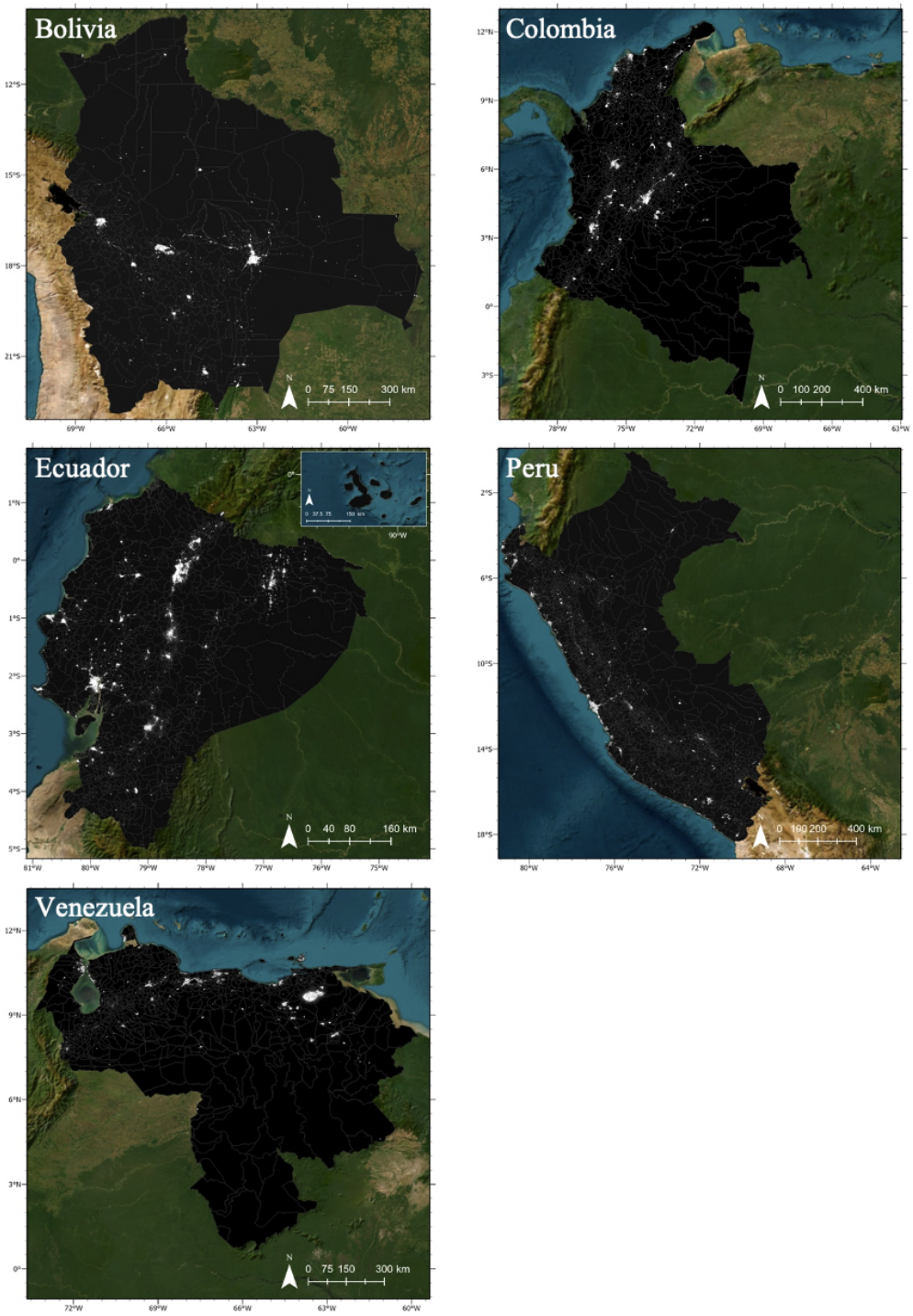
Using the latest VIIRS products has advantages. On the one hand, local studies using satellite light data often rely on older products, such as those from DMSP, publicly available from 1992 to 2013. VIIRS data outperform DMSP for mapping nighttime lights (Elvidge et al. 2013; Elvidge et al. 2017; Li et al. 2017; Levin et al. 2020). DMSP data have lower spatial resolution than VIIRS products, suffer from saturation, and have no onboard calibration.[14] Gibson and Boe-Gibson (2021) confirmed the DMSP data understate spatial inequality in, for example, the United States due to blurring and top-coding, while masked VIIRS products -version 2- perform better as proxies for local economic activity. Moreover, McCord and Rodriguez-Heredia (2022) found that nighttime lights strongly predict regional economic activity in seven countries, including Bolivia, Colombia, Ecuador, and Peru, using VIIRS data -version 2- from 2014 to 2019. Furthermore, I use version 2.1 as an upgrade of version 2, in which lit masks are updated, thus ensuring an improvement in the preprocessing. In this sense, this chapter gains reliability and precision using VIIRS data to obtain subnational outcomes.

Figure 4.2 shows the spatial distribution of light data across each country and the divisions by lowest geographic level. At first glance, there are notable differences in light data distribution by country.

[13] VIIRS products are available since April 2012. The annual products are processed from monthly raw data; thus, the 2012 annual data only considers the period from April 2012 to December 2012.

[14] These differences are motivated by sensor variations in spatial resolution, spectral response, point of spread function, overpass time at night, and wider radiance range of the VIIRS. Refer to Chapter 2 and Chapter 3.

Figure 4.2: Nighttime Lights by Andean Country, 2021



Source: Earth Observation Group, Payne Institute - Colorado School of Mines. Note: The maps are presented in alphabetical order.

In Bolivia, the intensity of light is spatially evident mainly around the city of Cochabamba (in central Bolivia in a valley in the Andes), in Santa Cruz de la Sierra (the largest city and principal industrial center in the country located on tropical lowlands, and one of the fastest growing cities in the region), and in its capital La Paz and the adjacent city El Alto on the Altiplano highlands.

Colombia also denotes multiple spots of light from main cities such as Bogotá (the capital and the most populated city in the country), Medellín (the second-largest city in the central region of the country, surrounded by the Andes Mountains), Cali (main urban and economic center in southwest Colombia, and third-largest city of the country), and Cartagena and Barranquilla in the north, Caribbean region, of the country.

Ecuador shows light concentration primarily through the Andean foothills from north to south (including its capital city, Quito), across its Pacific coastline including Guayaquil (principal economic capital and the port city of Ecuador), and in zones within the provinces of Orellana and Sucumbíos in the Amazon region in northeast Ecuador (territories containing multiple oil fields and relying primarily on exports of crude oil).

In Peru, light emissions are clearly visible throughout its vast coastline bordering the Pacific Ocean, with a significant concentration of light from Lima (the capital and largest city in the desert zone of the central coastal part) and dispersed spots, especially in the south and south central of the country.

Venezuela shows the most intensity of night lights in its northern area, mainly across its Andean region and the Coastal (Caribbean) Mountain range, in the northeast of Los Llanos region (a widely extended flat central depression), and surrounding the Maracaibo Basin located in the northwestern corner of Venezuela in Zulia state. Now, the highest intensity of light directly comes from natural gas flaring located primarily on oil fields in three out of eight oil-producing Venezuelan states: Anzoategui, Falcon, and Monagas.[15] As discussed in Chapter 2, light reflected from intensive gas flaring activity can be captured by satellite sensors. Treating gas flares as a measure of economic wealth could lead to misleading results on inequality. To reduce this concern, I excluded a total of 19,913 high light intensity pixels from Anzoategui, Falcon, and Monagas based on light distribution across their respective main urban cores. In particular, I use a state-specific

[15] As indicated in Chapter 3, Venezuela has eight oil-producing states: Anzoategui, Apure, Barinas, Delta Amacuro, Falcon, Guarico, Monagas, and Zulia); however, the highest intensity of light in the country come from pixels in Anzoategui, Falcon, and Monagas

cutoff value approach defined by the maximum urban core light intensity value unrelated to gas flaring across all the years, where values above the cutoff are flagged as light from gas flares in their respective state.[16]

Table 4.2 presents pixel-based summary statistics from 2012 to 2021. I extracted light data and population data, covering more than 222 million and 55.5 million observations, respectively. On average, only about 5 percent of the Andean region is entirely lit at night.[17] In comparison, about 45 percent of the territory seems to be populated. This context signals relative heterogeneity across the region.

Table 4.2: Summary Statistics (Pixel-Based), All Years

Country	Mean	Standard deviation	Minimum	Maximum	Observations
Bolivia					
Intensity of lights	2.0	14.0	0	100	52,240,480
Population (percent)	43.3	49.6	0	100	13,071,200
Population (count)	8.5	268.2	0	46,473	13,071,200
Colombia					
Intensity of lights	4.2	20.1	0	100	53,541,750
Population (percent)	45.7	49.8	0	100	13,419,930
Population (count)	35.7	748.3	0	92,873	13,419,930
Ecuador					
Intensity of lights	10.7	31.0	0	100	12,009,920
Population (percent)	62.8	48.3	0	100	3,019,180
Population (count)	53.5	644.5	0	46,214	3,019,180
Peru					
Intensity of lights	3.7	18.8	0	100	61,169,180
Population (percent)	46.3	49.9	0	100	15,315,360
Population (count)	20.5	465.1	0	63,485	15,315,360
Venezuela					
Intensity of lights	8.2	27.4	0	100	43,138,870
Population (percent)	37.9	48.5	0	100	10,802,420
Population (count)	27.9	478.5	0	62,924	10,802,420
Andean region					
Intensity of lights	4.7	21.1	0	100	222,100,200
Population (percent)	44.7	49.7	0	100	55,628,090
Population (count)	24.6	527.9	0	92,873	55,628,090

Keola et al. (2015) argued how agricultural activity emits marginal lights, if any; thus,

[16] I use the maximum value from urban pixels to avoid the elimination of urban pixels and reduce potential bias in the results.

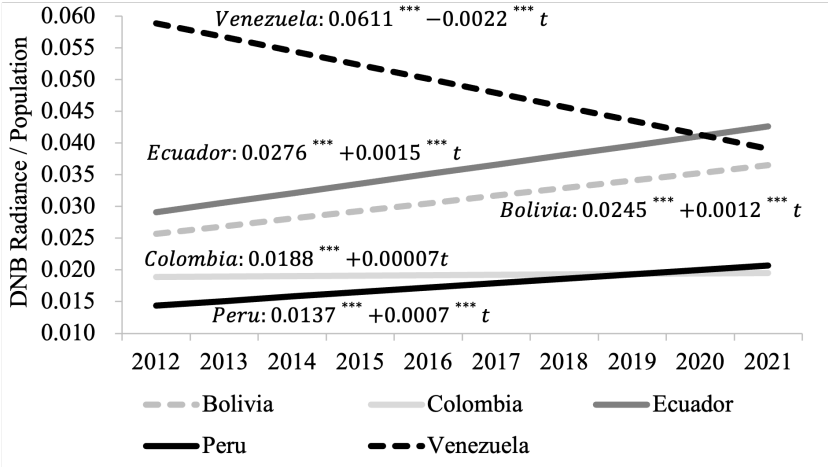
[17] The chapter assumes a minimum unit value to calculate the Theil index accounting for areas traditionally with very low intensity of light or no light at all captured by the sensor of satellites.

we can find vast land with low or no intensity of lights captured by satellite sensors. This situation may be especially true for Andean countries, highly rich in natural resources and where the agricultural sector has traditionally played an important economic and social role (Andrian and Manzano 2023).[18] An example of the disparity between the spatial distribution of light and population can be observed in Ecuador and Venezuela. Both countries have more lit areas than the Andean average (Ecuador with almost 11 percent and Venezuela with about 8 percent). Still, while nearly 38 percent of Venezuela's territory is inhabited, Ecuador has approximately 63 percent of its territory populated.

4.4 Results

Figure 4.3 shows how the intensity of light per capita evolved between 2012 and 2021 at the country level.[19] In general, the trends throughout the period are driven mainly by changes in the intensity of lights, except for Venezuela, where two main effects prevail, one due to light data and the other to population data.

Figure 4.3: Intensity of Light Per Capita (Linear Trend) by Andean Country



Source: Own calculations. Note: Significant at *10, ***1 percent, from estimates based on Ordinary Least Squares regressions, where t represents years.

[18] According to national sources, on average, between 2012 and 2021, the agricultural sector has represented 12.6 percent of GDP for Bolivia, 9.2 percent for Ecuador, 6.2 percent for Colombia, and 5.7 percent for Peru. Furthermore, the International Labour Organization (ILO) estimated agricultural employment represents 29.7 percent, 27.5 percent, 16.7 percent, and 27.8 percent of total employment in Bolivia, Ecuador, Colombia, and Peru, respectively. For details, see Andrian and Manzano (2023).

[19] I aggregate light data adding up the radiance outputs.

Bolivia, Ecuador, and Peru have been experiencing a significant upward trend. Ecuador shows the highest slope and reaches the highest intensity of light per inhabitant among the Andean countries. In particular, Ecuador averages an increase in the intensity of light per capita of 0.0015 per year, experiencing almost 50 percent more light per capita in 2019 than in 2012; then, the pandemic led to a sudden decrease of the indicator by 5.2 percent in 2020, followed by a recovery of 8.8 percent in 2021. Bolivia shows a slightly lower slope than Ecuador, averaging a 0.0012 increase per year. This country also had an increase of light per capita of about 50 percent from 2012 to 2019, which is more than 6 percent on average per year; however, the pandemic turned into a contraction of 2.2 percent and 1.8 percent in 2020 and 2021, respectively. The intensity of light per capita in Peru is growing at about 5.8 percent per year (60 percent of growth between 2012 and 2021). It is the only Andean country experiencing an increase during the pandemic. In the case of Colombia, the light per capita only increased nearly 6 percent between 2012 and 2021. In Venezuela, the indicator decreases parallel with the Venezuelan socioeconomic crisis, declining 30 percent between 2014 and 2021. However, since 2015 Venezuela has also experienced a massive international migration outflow still ongoing, which may even be attenuating the fall of the per capita indicator.[20]

A first glance at the Andean region as a whole should help identify dissimilarities at the country level. Moreover, the opposite trend found in Venezuela using light data and population data for recent years strengthens the idea of aggregate heterogeneities. For this, I use a two-level hierarchical structure, with the Andean region as the base level and the *Level* 1 of each country representing the underlying unit. Table 4.3 shows the one-stage Theil index decomposition accounting for all the countries in the region and testing the exclusion of each country from the sample.

[20]More than 7.2 million Venezuelans have fled the country as of March 2023. For details, see the Inter-Agency Coordination Platform for Refugees and Migrants from Venezuela (2023).

Table 4.3: One-Stage Theil Index Decomposition: Andean Region-Country

			Excluding									
	All		Bolivia		Colombia		Ecuador		Peru		Venezuela	
	2012	2021	2012	2021	2012	2021	2012	2021	2012	2021	2012	2021
Between-country	0.156	0.071	0.166	0.077	0.164	0.048	0.177	0.069	0.125	0.071	0.033	0.056
	(39.2)	*(23.5)*	*(39.5)*	*(23.3)*	*(38.0)*	*(15.0)*	*(42.8)*	*(23.3)*	*(32.4)*	*(21.3)*	*(17.4)*	*(31.8)*
Within-country	0.241	0.231	0.255	0.253	0.267	0.273	0.237	0.227	0.261	0.264	0.158	0.120
	(60.8)	*(76.5)*	*(60.5)*	*(76.7)*	*(62.0)*	*(85.0)*	*(57.2)*	*(76.7)*	*(67.6)*	*(78.7)*	*(82.6)*	*(68.2)*
Overall	0.397	0.302	0.421	0.331	0.431	0.321	0.414	0.295	0.385	0.335	0.191	0.176

Note: Contribution to the inequality in () in percentage.

The overall inequality for the Andean region using all the sample declined from 0.397 in 2012 to 0.302 in 2021. The between-country and within-country components reveal that the decrease in the overall inequality is due to the decline in the between-country component (from 0.156 to 0.071, respectively); in contrast, the within-country component was relatively stable. Furthermore, the relative importance of the between-country component decreased from about 39 percent in 2012 to 23.5 percent in 2021, while the within-country component gained weight on the inequality; thus, the overall wealth disparity in the region in 2021 is defined mainly by the inequality within each of the country and less by the differences among them.

An exclusion exercise over the sample finds similar patterns and somewhat similar inequality components, except when excluding Venezuela. For example, if Venezuela is considered as part of the sample and we exclude one by one the other countries, the overall inequality ranged from a minimum (maximum) of 0.385 (0.431) in 2012 to 0.295 (0.335) in 2021. However, suppose I exclude Venezuela and leave the remaining countries within the sample. In that case, the overall wealth inequality for the region dropped significantly, in absolute terms, to 0.191 in 2012 and 0.176 in 2021. Unlike the other scenarios, here, the between-country component gained instead of losing share (almost doubling in percentage contribution, from 17.4 percent to 31.8 percent, respectively).

These findings have at least three immediate implications. First, conducting a separate analysis by country is necessary to better understand inequality across the Andean region. Significant dissimilarities may be observed in recent years if Venezuela is not grouped with the other countries. Nevertheless, in any case, the within-country component still accounts for over 60 percent of total wealth inequality. Second, regional (aggregate) policy efforts to address wealth disparities have become more significant,

particularly when grouping Bolivia, Colombia, Ecuador, and Peru. However, including Venezuela highlights the need for increased attention to addressing within-country inequality in the region. Lastly, it is crucial to emphasize that further studies on inequality in the region should pay precise interest to the case of Venezuela.

Table 4.4 presents the decomposition results within each country. I adopt a four-level hierarchical structure for Bolivia, Ecuador, Peru, and Venezuela, enabling us to apply a three-stage nested Theil decomposition method. On the other hand, Colombia has a three-level hierarchical structure, so I follow a two-stage nested Theil decomposition method for this country.

Table 4.4: Multiple-Stage Nested Theil Index Decomposition by Andean Country

	Bolivia		Colombia		Ecuador		Peru		Venezuela	
	2012	2021	2012	2021	2012	2021	2012	2021	2012	2021
Between-0	0.009	0.003	0.095	0.050	0.145	0.124	0.005	0.014	0.228	0.308
	(2.2)	*(1.3)*	*(21.7)*	*(17.6)*	*(26.5)*	*(23.5)*	*(0.9)*	*(3.1)*	*(27.5)*	*(33.7)*
Between-1	0.049	0.021	0.063	0.047	0.130	0.127	0.087	0.055	0.112	0.161
	(12.5)	*(9.3)*	*(14.5)*	*(16.7)*	*(23.7)*	*(24.0)*	*(14.0)*	*(12.7)*	*(13.5)*	*(17.6)*
Between-2*	0.200	0.108	0.279	0.185	0.091	0.099	0.133	0.092	0.207	0.270
	(50.7)	*(48.0)*	*(63.8)*	*(65.7)*	*(16.6)*	*(18.7)*	*(21.6)*	*(21.2)*	*(24.9)*	*(29.5)*
Within-2	0.136	0.093			0.182	0.179	0.392	0.272	0.282	0.175
	(34.6)	*(41.3)*			*(33.1)*	*(33.8)*	*(63.5)*	*(63.0)*	*(34.0)*	*(19.2)*
Overall	0.394	0.224	0.437	0.282	0.549	0.530	0.617	0.433	0.829	0.913

Note: Contribution to the inequality in ( ) in percentage. *For Colombia, it corresponds to the Within-1 component from a two-stage nested Theil decomposition method based on *Level* 0-*Level* 1-*Level* 2.

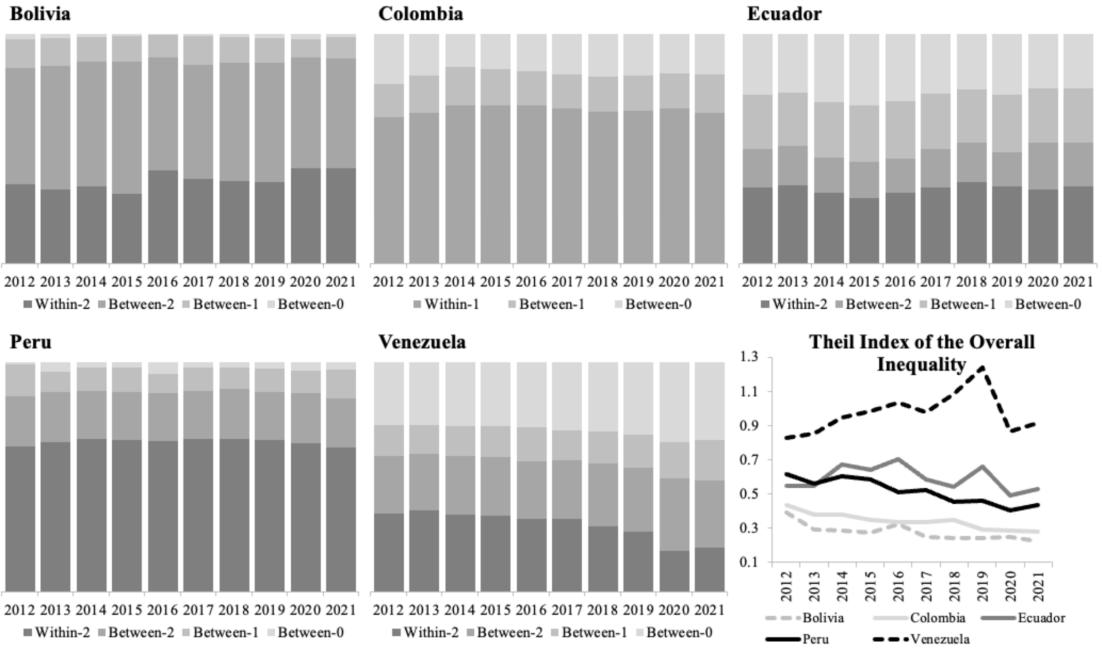
Bolivia exhibits the lowest overall Theil index, while Venezuela has the highest. Over the period from 2012 to 2021, Bolivia, Colombia, and Peru experienced a substantial decrease in wealth inequality. This reduction was observed across most components, except for Peru's between-0 component, which increased from 0.005 in 2012 to 0.014 in 2021.

In absolute terms, most of the overall reduction in inequality in Bolivia can be attributed to a decline in inequality between provinces. This indicates that the intensity of light has become more similar among provinces in 2021 compared to ten years ago. On the other hand, Colombia and Peru primarily witnessed a decrease in inequality within municipalities and districts, respectively. Ecuador has little change in overall inequality; this is also true for its components. In contrast, Venezuela stands out as the only Andean country where overall inequality has increased, primarily driven by

changes in inequality between all geographic subgroups. However, inequality within municipalities has been actually decreasing in Venezuela.

Figure 4.4 illustrates the variations in the contributions of different components from 2012 to 2021. Wealth disparities associated with the lowest-level administrative divisions have significantly contributed to overall inequality in each Andean country.

Figure 4.4: Relative Importance of the Spatial Dimension on Overall Inequality by Country



Source: Own calculations.

Approximately 90 percent of total inequality in Bolivia can be attributed to within and between components related to provinces in 2021. Furthermore, the decrease in inequality in Bolivia has amplified the relative importance of inequalities within provinces, accounting for 41.3 percent. Consequently, it is critical for Bolivia to implement policy efforts promoting wealth equality at the provincial level.

In Peru, the within and between district components in 2021 constitute about 84.2 percent of total inequality, representing 63 percent and 21.2 percent, respectively. The decrease in inequality in Peru has had little impact on the contribution of these components. Similarly to Bolivia, focusing on the district level may be a suitable local

scale for policy interventions to address overall inequality in Peru, specifically by reducing disparities within districts.

Inequality has decreased across all spatial components in Colombia, although the disparities within municipalities are slightly gaining relative prominence, representing 65.7 percent in 2021. Therefore, implementing local programs to combat inequality at the municipal level would be a proper approach in Colombia.

In Ecuador and Venezuela, the inequality components exhibit a more balanced contribution to overall inequality. In particular, these countries require special attention among Andean countries. The results indicate that Ecuador has not undergone a significant reduction in inequality over a decade, whereas the findings confirm an increase in overall inequality in Venezuela. In Ecuador, the disparity within cantons has been the most significant contributor to total inequality, accounting for approximately one-third of the overall inequality on average. Conversely, in Venezuela, the differences among regions, *Level* 0, and between municipalities, *Level* 2, traditionally represent the major drivers of overall inequality (which may signal a different aggregate degree of territorial economic development than the rest of Andean countries). Therefore, it seems plausible that it will be necessary to implement local policies and national programs encompassing multiple spatial levels in both countries to address the prevailing inequality challenges.

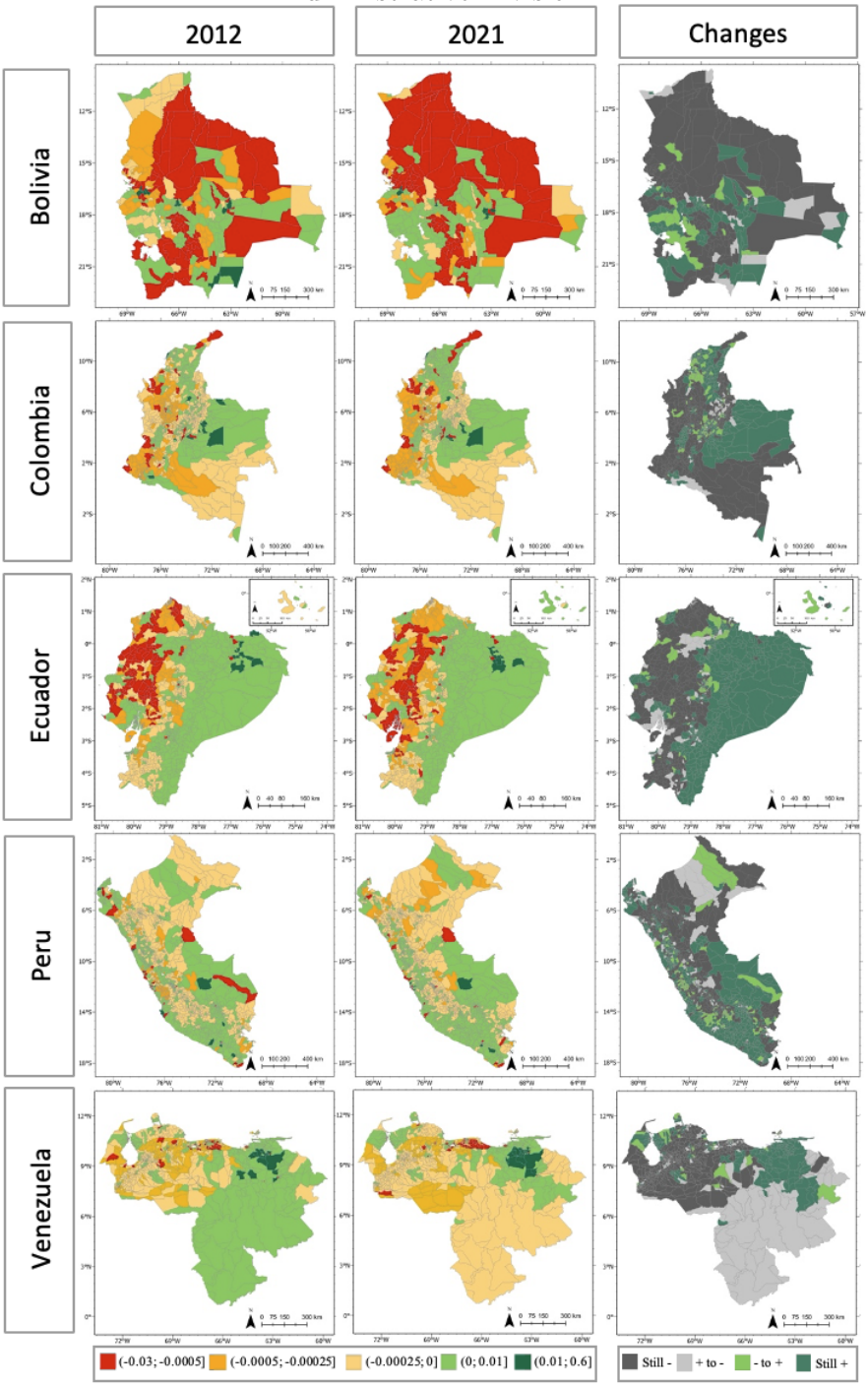
Finally, to examine the presence of spatial heterogeneities, Figure 4.5 exhibits the raw contributions to the Theil index in 2012, 2021, and the change between these two years, categorized by country at the lowest-level administrative division.

In this chapter, the index captures the degree to which the distribution of light data among different groups differs from the distribution of population data among those same groups. In this sense, groups with higher proportions of the intensity of light compared to their population shares contribute positively to the Theil index; in contrast, those with lower shares of the intensity of light relative to their population shares contribute negatively.

In the columns of each year, red and dark green areas represent spatial locations where the relative proportions are strongly negative and strongly positive, respectively, while yellowish and light green areas are contributions closer to zero. In the column of changes, green areas must draw our main attention.

In the past decade, approximately one-third of Bolivia and Colombia, and around

Figure 4.5: Contributions to the Theil Index by Andean Country: Lowest-Level Administrative Division



Source: Own calculations.

40 percent of Peru, have continued to show positive contributions to the Theil index, indicating persistent disparities where light data surpasses population data. These areas represent "structural" challenges in terms of wealth inequality. Bolivia and Peru experienced a similar increase/decrease in areas denoting wealth inequality (around 8 percent and 10 percent, respectively). The situation differs in Colombia, where 10 percent of its territory faced newly emerged areas characterized by wealth inequality or recent inequality spots (light green in the last column), and only 6 percent resulted in newly emerged areas leaving wealth inequality (i.e., a net increase of 4 percent of its territory facing wealth inequality). Despite these variations, all three countries have experienced a significant reduction in overall inequality.[21] This suggests that there has been a general decrease in wealth inequality across areas where inequality has traditionally prevailed.

Ecuador's overall inequality has shown minimal change. Notably, 43 percent of its territory still experienced wealth inequality in 2021 compared to 2012. There was also a net increase of 3 percent in the territory with wealth inequality, with 9 percent of newly emerged parishes demonstrating wealth inequality compared to 6 percent of parishes where inequality had "ceased." This context implies that wealth inequality is spreading across the territory. It highlights the need to reinforce or reassess policy efforts to address wealth disparities in traditionally unequal areas (darker green) and focus on taking action in new areas that are now exhibiting inequalities (light green).

Here, again, Venezuela differs significantly from the rest of the countries. In 2021, wealth inequality appeared to be highly concentrated in approximately 30 percent of its territory. This concentration can be attributed to persistent wealth disparities in 26 percent of the parishes, as well as the emergence of new parishes (4 percent) where inequality has manifested. The stark spatial concentration of wealth inequality suggests significant variations in economic development across different regions of the country. Furthermore, Venezuela is the only Andean country where the Theil index has increased, indicating a net worsening of wealth inequality. This underscores the urgency for implementing national and local policy actions that address these inequalities and work towards reducing overall wealth disparities within the country.

[21] For details, see Table 4.4.

4.5 Conclusion

In recent years, insights into the extent and spatial nature of inequality across countries have increasingly relied on novel approaches to gather comprehensive data and analyze patterns on a local and global scale. One such approach involves applying remote sensing techniques. This chapter uses satellite-recorded nighttime lights and gridded population datasets from 2012 to 2021 to shed light on the appropriate local scale of action for initiatives to tackle inequality across each country of the Andean region. In this sense, I follow multiple-stage nested Theil decomposition method using night light emissions as a proxy of economic wealth.

The study confirms a reduction in overall inequality for the Andean region between 2012 and 2021 (primarily driven by a decline in between-country inequality) and an increase in the relative importance of within-country inequality. This result emphasizes the need for regional policy efforts to address wealth disparities while recognizing the unique challenges posed by each country. Moreover, the chapter also verifies the urgency of addressing wealth inequality in Venezuela, suggesting focusing on this country for further research on inequality.

The main results also reveal spatial heterogeneities by country. Bolivia, Colombia, and Peru experienced a substantial decline in wealth inequality over the past decade. Bolivia's reduction in inequality can be attributed to the decline in disparities between provinces, while Colombia and Peru witnessed decreases within municipalities and districts, respectively. They also experienced a significant increase in the relative contribution of their within-inequality component, emphasizing the need for targeted policy interventions at the provincial level in Bolivia, municipal level in Colombia, and district level in Peru.

On the other hand, Ecuador has shown a minimal change in overall inequality, and Venezuela stands out as the only country in the Andean region where inequality has increased, primarily driven by changes in the disparity between all geographic subgroups. Ecuador and Venezuela exhibit more balanced contributions of their components to their inequality results, although Ecuador's cantons and Venezuela's regions play crucial roles.

The reduction of inequality must be carried out by different mechanisms operating on the spatial scale. Policy efforts at the lowest division must have a central role in reducing

wealth inequality in the Andean countries, particularly in Bolivia, Colombia, and Peru, while, in Ecuador and Venezuela prevails a major urgency for the implementation and reinforcement of national and local policy actions that address these inequalities and work towards reducing overall wealth disparities. The main findings underscore the persistent challenges in wealth inequality and highlight the potential for targeted interventions and policy reforms at different spatial scales to mitigate those disparities and foster more equitable societies in the Andean region.

Bibliography

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

Akita, T. (2003). Decomposing regional income inequality in China and Indonesia using two-stage nested Theil decomposition method. *The Annals of Regional Science*, *37*, 55–77.

Alstadsæter, A., Johannesen, N., & Zucman, G. (2019). Tax evasion and inequality. *American Economic Review*, *109*(6), 2073–2103.

Amundsen, I. (2014). Drowning in Oil: Angola's Institutions and the "Resource Curse". *Comparative politics*, *46*(2), 169–189.

Andersen, L. E., Guzmán, G., & Garbasevshi, O. (2023, January). *Electricity Consumption, Nightlights, and Local Development in Bolivia*. https://sdsnbolivia.org/en/electricity-consumption-nightlights-and-local-development-in-bolivia/

Andrade-Núñez, M. J., & Aide, T. M. (2020). Using nighttime lights to assess infrastructure expansion within and around protected areas in South America. *Environmental Research Communications*, *2*(2), 021002.

Andreano, M. S., Benedetti, R., Piersimoni, F., & Savio, G. (2021). Mapping poverty of Latin American and Caribbean countries from heaven through night-light satellite images. *Social Indicators Research*, *156*(2-3), 533–562.

Andrian, L. G., & Manzano, O. (2023). Nuevos horizontes de transformación productiva en la Región Andina. *Monografía del BID*, *1022*.

Arderne, C., Zorn, C., Nicolas, C., & Koks, E. (2020). Predictive mapping of the global power system using open data. *Scientific data*, *7*(1), 1–12.

Ayala, L., Pérez, A., & Prieto-Alaiz, M. (2022). The Impact of Different Data Sources on the Level and Structure of Income Inequality. *SERIEs*, *13*(3), 583–611.

Bazilian, M., Onyeji, I., Aqrawi, P.-K., Sovacool, B. K., Ofori, E., Kammen, D. M., & Van de Graaf, T. (2013). Oil, energy poverty and resource dependence in West Africa. *Journal of Energy & Natural Resources Law*, *31*(1), 33–53.

Bird, K. (2013). The intergenerational transmission of poverty: An overview. *Chronic poverty: Concepts, causes and policy*, 60–84.

Bluhm, R., & Krause, M. (2018). *Top Lights-Bright Cities and their Contribution to Economic Development* (tech. rep.). CESifo Working Paper No. 7411.

Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica*, *47*(4), 901–920.

Bruederle, A., & Hodler, R. (2018). Nighttime lights as a proxy for human development at the local level. *PloS one*, *13*(9).

Burchfield, M., Overman, H. G., Puga, D., & Turner, M. A. (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics*, *121*(2), 587–633.

Burkhauser, R. V., Feng, S., Jenkins, S. P., & Larrimore, J. (2012). Recent trends in top income shares in the United States: Reconciling estimates from March CPS and IRS tax return data. *Review of Economics and Statistics*, *94*(2), 371–388.

Bustos, M. F. A., Hall, O., Niedomysl, T., & Ernstson, U. (2020). A pixel level evaluation of five multitemporal global gridded population datasets: a case study in Sweden, 1990–2015. *Population and environment*, 1–23.

Cabrera-Barona, P. F., Bayón, M., Durán, G., Bonilla, A., & Mejía, V. (2020). Generating and mapping Amazonian urban regions using a geospatial approach. *ISPRS International Journal of Geo-Information*, *9*(7), 453.

Carr, M. D., & Wiemers, E. E. (2018). New evidence on earnings volatility in survey and administrative data. *AEA Papers and Proceedings*, *108*, 287–291.

Castañeda, A., Doan, D., Newhouse, D., Nguyen, M. C., Uematsu, H., Azevedo, J. P., et al. (2018). A new profile of the global poor. *World Development*, *101*, 250–267.

Castilleja, L. (2020). La clase media andina frente al shock del Covid-19. *Documento para Discusión IDB-DP-00774*.

Cerra, M. V. (2016). *Inflation and the Black Market Exchange Rate in a Repressed Market: A Model of Venezuela*. International Monetary Fund.

Chen, X. (2016). Addressing measurement error bias in GDP with nighttime lights and an application to infant mortality with Chinese county data. *Sociological Methodology*, *46*(1), 319–344.

Conceição, P., & Ferreira, P. (2000). *The young person's guide to the Theil index: Suggesting intuitive interpretations and exploring analytical applications* (UTIP Working Paper No. 14). University of Texas at Austin.

Courtemanche, C., Denteh, A., & Tchernis, R. (2019). Estimating the associations between SNAP and food insecurity, obesity, and food purchases with imperfect administrative measures of participation. *Southern Economic Journal*, *86*(1), 202–228.

Croft, T. A. (1978). Nighttime images of the earth from space. *Scientific American*, *239*(1), 86–101.

Dahl, M., DeLeire, T., & Schwabish, J. A. (2011). Estimates of year-to-year volatility in earnings and in household incomes from administrative, survey, and matched data. *Journal of Human Resources*, *46*(4), 750–774.

Dai, Z., Hu, Y., & Zhao, G. (2017). The suitability of different nighttime light data for GDP estimation at different spatial scales and regional levels. *Sustainability*, *9*(2), 305.

De la Cruz, R., Manzano, O., & Loterszpil, M. (2020). Cómo acelerar el crecimiento económico y fortalecer la clase media: América Latina. *Monografía del BID*, *782*.

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

Deaton, A. (2016). Measuring and understanding behavior, welfare, and poverty. *American Economic Review*, *106*(6), 1221–1243.

Debbich, M. (2019). *Assessing Oil and Non-Oil GDP Growth from Space: An Application to Yemen 2012-17*. International Monetary Fund.

Denis, M. (2021). Venezuela: Confrontación política, límite de la polarización y crisis de la representación. *Revista Cuerpo y Territorio*.

Do, Q.-T., Shapiro, J. N., Elvidge, C. D., Abdel Jelil, M., Ahn, D. P., Baugh, K., Hansen-Lewis, J., & Zhizhin, M. (2017). How much oil is the Islamic state group producing? Evidence from remote sensing. *Evidence from Remote Sensing (October 31, 2017). World Bank Policy Research Working Paper*, (8231).

Doll, C. N., Muller, J.-P., & Morley, J. G. (2006). Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics*, *57*(1), 75–92.

Earth Observation Group. (2021, October). *Payne Institute for Public Policy - Colorado School of Mines* [Accessed on October 14, 2021]. https://payneinstitute.mines.edu/eog/nighttime-lights/

Earth Observation Group. (2022, June). *Payne Institute for Public Policy - Colorado School of Mines* [Accessed on June 7, 2022]. https://payneinstitute.mines.edu/eog/nighttime-lights/

Economic Commission for Latin America and the Caribbean, ECLAC, (2018). *Ruralidad, hambre y pobreza en América Latina y el Caribe* (tech. rep. No. LC/TS.2018/119). ECLAC. Santiago.

Economic Commission for Latin America and the Caribbean's Statistical Database. (2022, January). Eclacstat [Accessed on January 12, 2023]. United Nations. https://statistics.cepal.org/portal/cepalstat/dashboard.html?lang=en

Elbers, C., Lanjouw, P., Mistiaen, J. A., & Özler, B. (2008). Reinterpreting between-group inequality. *The Journal of Economic Inequality*, *6*, 231–245.

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R., & Davis, C. W. (1997). Relation between satellite observed visible-near infrared emissions, population, economic activity, and electric power consumption. *International Journal of Remote Sensing*, *18*(6), 1373–1379.

Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F.-C., & Ghosh, T. (2017). VIIRS night-time lights. *International Journal of Remote Sensing*, *38*(21), 5860–5879.

Elvidge, C. D., Baugh, K. E., Zhizhin, M., & Hsu, F.-C. (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proceedings of the Asia-Pacific Advanced Network*, *35*(62).

Elvidge, C. D., 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*(5), 922.

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

ENCOVI. (2021). *National Survey of Living Conditions, Venezuela*. https://www.proyectoencovi.com/encovi-2021

ENCOVI. (2019-2020). *National Survey of Living Conditions, Venezuela*. https://www.proyectoencovi.com/informe-interactivo-2019

Ferreira, T. (2018). *Using satellite data to track socio-economic outcomes: a case study of Namibia* (tech. rep.). Stellenbosch University, Department of Economics.

Floerkemeier, H., Spatafora, N., & Venables, A. (2021). Regional disparities, growth, and inclusiveness. *International Monetary Fund (IMF) Working Paper No. 38*.

Foster, J. (1983). Observations of the Earth using nighttime visible imagery. *Optical Engineering for Cold Environments*, *414*, 187–193.

Galbraith, J. K. (2019). Sparse, inconsistent and unreliable: Tax records and the world inequality report 2018. *Development and Change*, *50*(2), 329–346.

Galimberti, J. K., Pichler, S., & Pleninger, R. (2023). Measuring inequality using geospatial data. *The World Bank Economic Review*, *37*(4), 549–569.

Getie, E. M. (2020). Poverty of Energy and Its Impact on Living Standards in Ethiopia. *Journal of Electrical and Computer Engineering*, *2020*.

Ghosh, T., L Powell, R., D Elvidge, C., E Baugh, K., C Sutton, P., & Anderson, S. (2010). Shedding light on the global distribution of economic activity. *The Open Geography Journal*, *3*(1).

Gibson, J., & Boe-Gibson, G. (2021). Nighttime lights and county-level economic activity in the United States: 2001 to 2019. *Remote Sensing*, *13*(14), 2741.

Gilliland, J. A., Shah, T. I., Clark, A., Sibbald, S., & Seabrook, J. A. (2019). A geospatial approach to understanding inequalities in accessibility to primary care among vulnerable populations. *PloS One*, *14*(1), e0210113.

González Oquendo, L. J. (2019). Colapso eléctrico y colapso gerencial en Venezuela. *Revista Venezolana de Gerencia*.

González-Eguino, M. (2015). Energy poverty: An overview. *Renewable and sustainable energy reviews*, *47*, 377–385.

Guevara Baro, M. (2020). El Colapso Eléctrico de Venezuela y los Desafíos para Superarlo. *Debates IESA*, Instituto De Estudios Superiores De Administración.

Haithcoat, T. L., Avery, E. E., Bowers, K. A., Hammer, R. D., & Shyu, C.-R. (2021). Income inequality and health: Expanding our understanding of state-level effects

by using a geospatial big data approach. *Social Science Computer Review*, *39*(4), 543–561.

Halff, A., Monaldi, F., Palacios, L., & Santos, M. A. (2017). Apocalypse Now: Venezuela, Oil, and Reconstruction.

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

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

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

Inter-Agency Coordination Platform for Refugees and Migrants from Venezuela. (2022). *R4V (website)* [Accessed on October 20, 2022]. https://www.r4v.info/

Inter-Agency Coordination Platform for Refugees and Migrants from Venezuela. (2023). *R4V (website)* [Accessed on March 15, 2023]. https://www.r4v.info/

International Energy Agency, International Renewable Energy Agency, UN Department of Economic and Social Affairs, World Bank, & World Health Organization. (2022). Tracking SDG 7: The Energy Progress Report.

Jagnani, M., & Khanna, G. (2020). The effects of elite public colleges on primary and secondary schooling markets in India. *Journal of Development Economics*, *146*, 102512.

Jeswani, R., Kulshrestha, A., Gupta, P. K., & Srivastav, S. (2019). Evaluation of the consistency of DMSP-OLS and SNPP-VIIRS Night-time Light Datasets. *J. Geomat*, *13*, 98–105.

Keola, S., Andersson, M., & Hall, O. (2015). Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. *World Development*, *66*, 322–334.

Khan, M. S., & Siddique, A. B. (2021). Spatial analysis of regional and income inequality in the United States. *Economies*, *9*(4), 159.

Khandker, S. R., Samad, H. A., Ali, R., & Barnes, D. F. (2014). Who benefits most from rural electrification? Evidence in India. *The Energy Journal*, *35*(2).

Kharas, H., Di Nucci, C., Hamel, K., & Tong, B. (2020). To move the needle on ending extreme poverty, focus on rural areas.

Kurmanaev, A. (2019). Venezuela's collapse is the worst outside of war in decades, economists say. *New York Times*.

Lanjouw, P. F., & Tarp, F. (2021). Poverty, vulnerability and Covid-19: Introduction and overview. *Review of development economics*, *25*(4), 1797–1802.

Lee, K., Brewer, E., Christiano, C., Meyo, F., Miguel, E., Podolsky, M., Rosa, J., & Wolfram, C. (2016). Electrification for "under grid" households in rural Kenya. *Development Engineering*, *1*, 26–35.

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

Levin, N., Kyba, C. C., Zhang, Q., de Miguel, A. S., Román, M. O., Li, X., Portnov, B. A., Molthan, A. L., Jechow, A., Miller, S. D., et al. (2020). Remote sensing of night lights: A review and an outlook for the future. *Remote Sensing of Environment*, *237*, 111443.

Li, X., Li, D., 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. *International Journal of Remote Sensing*, *38*(21), 5934–5951.

Li, X., Zhou, Y., Zhao, M., & Zhao, X. (2020). A harmonized global nighttime light dataset 1992–2018. *Scientific Data*, *7*(1), 1–9.

Llambí, L., Arias, E., & Briceño, G. (1994). *Pequeña producción agrícola y pobreza rural en Venezuela* (tech. rep.). Instituto Venezolano de Investigaciones Científicas, Caracas (Venezuela ...

Loayza, N., & Rigolini, J. (2016). The local impact of mining on poverty and inequality: evidence from the commodity boom in Peru. *World Development*, *84*, 219–234.

Lynn, P., Jäckle, A., Jenkins, S. P., & Sala, E. (2012). The impact of questioning method on measurement error in panel survey measures of benefit receipt: Evidence from a validation study. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *175*(1), 289–308.

Ma, J., Guo, J., Ahmad, S., Li, Z., & Hong, J. (2020). Constructing a New Inter-Calibration Method for DMSP-OLS and NPP-VIIRS Nighttime Light. *Remote Sensing*, *12*(6).

Maldonado, L. (2022). Lighting-up the economic activity of oil-producing regions: A remote sensing application. *Remote Sensing Applications: Society and Environment*, *26*, 100722.

Maldonado, L. (2023). Living in darkness: Rural poverty in Venezuela. *Journal of Applied Economics*, *26*(1), 2168464.

Maldonado, L., & Olivo, V. (2022). Is Venezuela still an upper-middle-income country? estimating the GNI per capita for 2015–2021. *IADB. December*.

Manzano, O., Monaldi, F., & Sturzenegger, F. (2008). The political economy of oil production in Latin America [with comments]. *Economía*, *9*(1), 59–103.

Márquez, G. (1994). Pobreza y políticas sociales en Venezuela. *Contribuciones*, *1*, 137–174.

Martel, A. (1995). La pobreza rural en Venezuela. *Revista Venezolana de Analísis de Coyuntura*, *1*(1).

Martinez, L. R. (2022). How much should we trust the dictator's GDP growth estimates? *Journal of Political Economy*, *130*(10), 2731–2769.

McCord, G. C., & Rodriguez-Heredia, M. (2022). Nightlights and subnational economic activity: Estimating departmental GDP in Paraguay. *Remote Sensing*, *14*(5), 1150.

Mejía Juárez, V. (2020). Procesos de urbanización y morfología urbana en Ecuador: La evolución de los usos del suelo a la luz de la imagen satelital nocturna de la Tierra 1992-2019.

Mendoza Jr, C. B., Cayonte, D. D. D., Leabres, M. S., & Manaligod, L. R. A. (2019). Understanding multidimensional energy poverty in the Philippines. *Energy Policy*, *133*, 110886.

Meyer, B. D., & Mittag, N. (2021). Combining administrative and survey data to improve income measurement. *Administrative Records for Survey Methodology*, 297–322.

Min, B. K.-H. (2010). *Democracy and light: public service provision in the developing world*. University of California, Los Angeles.

Mirza, M. U., Xu, C., Bavel, B. v., Van Nes, E. H., & Scheffer, M. (2021). Global inequality remotely sensed. *Proceedings of the National Academy of Sciences*, *118*(18).

Morduch, J., & Sicular, T. (2002). Rethinking inequality decomposition, with evidence from rural China. *The Economic Journal*, *112*(476), 93–106.

Moss, T., Bazilian, M., Blimpo, M., Culver, L., Kincer, J., Mahadavan, M., Modi, V., Muhwezi, B., Mutiso, R., Sivaram, V., et al. (2020). The Modern Energy Minimum: The case for a new global electricity consumption threshold. *Energy for Growth Hub*, *30*.

Mveyange, A. (2015). *Night Lights and Regional Income Inequality in Africa* (tech. rep. No. 2015/085). WIDER Working Paper.

Njiru, C. W., & Letema, S. C. (2018). Energy poverty and its implication on standard of living in Kirinyaga, Kenya. *Journal of Energy*, *2018*.

Nordhaus, W. D., & Chen, X. (2012). Improved estimates of using luminosity as a proxy for economic statistics: new results and estimates of precision.

Oak Ridge National Laboratory, ORNL. (2021, March). *LandScan Global Population Database. Oak Ridge, TN: Oak Ridge National Laboratory.*

Oak Ridge National Laboratory, ORNL. (2023, February). *LandScan Global Population Database. Oak Ridge, TN: Oak Ridge National Laboratory.*

Olinto, P., Beegle, K., Sobrado, C., Uematsu, H., et al. (2013). The state of the poor: Where are the poor, where is extreme poverty harder to end, and what is the current profile of the world's poor. *Economic Premise*, *125*(2), 1–8.

Olivo, V., & Saboin, J. L. (2020). Venezuela's Lagged Price Adjustment: Inflationary Pass-through, Consumption and Distributional Impacts, and (Potential) Policy Implications.

Paredes, D., Iturra, V., & Lufin, M. (2012). *A further step to understand income inequality in Chile: A decomposition using three-stages nested Theil index decomposition method* (tech. rep. No. 25). Universidad Católica del Norte, Chile, Department of Economics.

Parés-Ramos, I. K., Álvarez-Berríos, N. L., & Aide, T. M. (2013). Mapping urbanization dynamics in major cities of Colombia, Ecuador, Perú, and Bolivia using night-time satellite imagery. *Land*, *2*(1), 37–59.

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

Prensa Aula Abierta. (2020). *Racionamientos, apagones o servicio inexistente: Así se ve la crisis eléctrica en Venezuela* [Accessed on August 21, 2022]. https://aulaabiertavenezuela.org/index.php/2020/05/06/racionamientos-apagones-o-servicio-inexistente-asi-se-ve-la-crisis-electrica-en-venezuela/

Puttanapong, N., Luenam, A., & Jongwattanakul, P. (2022). Spatial analysis of inequality in Thailand: Applications of satellite data and spatial statistics/econometrics. *Sustainability*, *14*(7), 3946.

Rabiei-Dastjerdi, H., & Matthews, S. A. (2021). Who gets what, where, and how much? Composite index of spatial inequality for small areas in Tehran. *Regional Science Policy & Practice*, *13*(1), 191–205.

Ritchie, H., & Roser, M. (2019). Access to energy. *Our World in Data*.

Riutort, M. (1999). Las causas de la pobreza en Venezuela. *Revista temas de coyuntura*, (40).

Rodríguez-Pose, A., & Hardy, D. (2015). Addressing poverty and inequality in the rural economy from a global perspective. *Applied Geography*, *61*, 11–23.

Ross, M. L. (2001). Does oil hinder democracy? *World politics*, *53*(3), 325–361.

Sabatini, C., & Patterson, W. (2021). Reforming Venezuela's electricity sector. *US and the Americas Programme. Environment and Society Programme. Chatam House (Research Paper)*.

Sachs, J., Lafortune, G., Kroll, C., Fuller, G., & Woelm, F. (2022). *From Crisis to Sustainable Development: the SDGs as Roadmap to 2030 and Beyond* (Sustainable Development Report 2022.). Cambridge University Press. Cambridge.

Sánchez de Miguel, A., Bennie, J., Rosenfeld, E., Dzurjak, S., & Gaston, K. J. (2021). First estimation of global trends in nocturnal power emissions reveals acceleration of light pollution. *Remote Sensing*, *13*(16), 3311.

Sánchez de Miguel, A., Zamorano, J., Castaño, J. G., & Pascual, S. (2014). Evolution of the energy consumed by street lighting in Spain estimated with DMSP-OLS data. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *139*, 109–117.

Seminario, B., & Palomino, L. (2022). *Estimación del PIB a nivel subnacional utilizando datos satelitales de luminosidad: Perú, 1993-2018*. Universidad del Pacífico.

Shorrocks, A., & Wan, G. (2005). Spatial decomposition of inequality. *Journal of Economic Geography*, *5*(1), 59–81.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, *48*(3), 613–625.

Sinha, A., Balsalobre-Lorente, D., Zafar, M. W., & Saleem, M. M. (2022). Analyzing global inequality in access to energy: Developing policy framework by inequality decomposition. *Journal of Environmental Management*, *304*, 114299.

Smith, B., & Wills, S. (2018). Left in the dark? Oil and rural poverty. *Journal of the Association of Environmental and Resource Economists*, *5*(4), 865–904.

Spagnoletti, B., & O'Callaghan, T. (2013). Let there be light: A multi-actor approach to alleviating energy poverty in Asia. *Energy policy*, *63*, 738–746.

Stampini, M., Robles, M., Sáenz, M., Ibarrarán, P., & Medellín, N. (2015). *Pobreza, vulnerabilidad y la clase media en América Latina* (tech. rep.). IDB Working Paper Series.

Sullivan, W. T. (1989). A 10 km resolution image of the entire night-time Earth based on cloud-free satellite photographs in the 400–1100 nm band. *Remote Sensing*, *10*(1), 1–5.

Sutton, P. C. (2003). A scale-adjusted measure of "urban sprawl" using nighttime satellite imagery. *Remote sensing of environment*, *86*(3), 353–369.

Sutton, P. C., & Costanza, R. (2002). Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. *Ecological Economics*, *41*(3), 509–527.

Theil, H. (1967). *Economics and information theory*. North-Holland Pub. Co.; Rand McNally.

Venezuelan Observatory of Public Services. (2021). *Nuevos Resultados de la Encuesta de Percepción Ciudadana sobre los Servicios Públicos. Agosto-Septiembre 2021. Boletín Informativo del OVSP, 28.* [Accessed on September 15, 2022]. https://www.observatoriovsp.org/electricidad/

Venezuelan Observatory of Social Conflict. (2020). *Social Conflict in Venezuela in 2019* [Accessed on November 3, 2021]. https://www.observatoriodeconflictos.org.ve/tendencias-de-la-conflitividad/conflictividad-social-en-venezuela-en-2019

Vera, L. (2017). In search of stabilization and recovery: macro policy and reforms in Venezuela. *Journal of Post Keynesian Economics*, *40*(1), 9–26.

Wang, X., Rafa, M., Moyer, J. D., Li, J., Scheer, J., & Sutton, P. (2019). Estimation and mapping of sub-national GDP in Uganda Using NPP-VIIRS imagery. *Remote Sensing*, *11*(2), 163.

Weidmann, N. B., & Schutte, S. (2017). Using night light emissions for the prediction of local wealth. *Journal of Peace Research*, *54*(2), 125–140.

Weidmann, N. B., & Theunissen, G. (2021). Estimating local inequality from nighttime lights. *Remote Sensing*, *13*(22), 4624.

Welch, R. (1980). Monitoring urban population and energy utilization patterns from satellite data. *Remote sensing of Environment*, *9*(1), 1–9.

World Bank. (2018). Poverty and shared prosperity 2018: piecing together the poverty puzzle.

Wu, D., Yuan, L., Li, R., & Li, J. (2018). Decomposing inequality in research funding by university-institute sub-group: A three-stage nested Theil index. *Journal of Informetrics*, *12*(4), 1312–1326.

Wu, K., & Wang, X. (2019). Aligning pixel values of DMSP and VIIRS nighttime light images to evaluate urban dynamics. *Remote Sensing*, *11*(12), 1463.

Yin, X., Li, P., Feng, Z., Yang, Y., You, Z., & Xiao, C. (2021). Which gridded population data product is better? Evidences from mainland Southeast Asia (MSEA). *ISPRS International Journal of Geo-Information*, *10*(10), 681.

Zhang, L., Li, X., & Chen, F. (2020). Spatiotemporal analysis of Venezuela's nighttime light during the socioeconomic crisis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2396–2408.

Zhao, M., Zhou, Y., Li, X., Zhou, C., Cheng, W., Li, M., & Huang, K. (2019). Building a Series of Consistent Night-Time Light Data (1992–2018) in Southeast Asia by Integrating DMSP-OLS and NPP-VIIRS. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(3), 1843–1856.

Appendix A

Supplemental Material: Chapter 2

A.1 Regions in the Sample

Table A.1: Regions: Selected Sample

Brazil	Canada	China	Colombia	India	Mexico
Acre	Alberta*	Anhui	Amazonas	Andaman and Nicobar	Aguascalientes
Alagoas	British Columbia*	Beijing	Antioquia*	Andhra Pradesh*	Baja California
Amapá	Manitoba*	Chongqing*	Arauca*	Arunachal Pradesh	Baja California Sur
Amazonas	New Brunswick	Fujian	Atlántico*	Assam*	Campeche
Bahia	Newfoundland and Labrador*	Gansu*	Bolívar*	Bihar	Chiapas*
Ceará	Northwest Territories	Guangdong	Boyacá*	Chandigarh	Chihuahua
Distrito Federal	Nova Scotia	Guangxi	Caldas	Chhattisgarh	Coahuila
Espírito Santo	Nunavut	Guizhou	Caquetá	Goa	Colima
Goiás	Ontario*	Hainan	Casanare*	Gujarat*	Distrito Federal
Maranhão	Prince Edward Island	Hebei*	Cauca*	Haryana	Durango
Mato Grosso	Québec	Heilongjiang*	Cesar*	Himachal Pradesh	Guanajuato
Mato Grosso do Sul	Saskatchewan*	Henan	Chocó	Jammu and Kashmir	Guerrero
Minas Gerais	Yukon	Hubei	Córdoba*	Jharkhand	Hidalgo
Pará		Hunan	Cundinamarca*	Karnataka	Jalisco
Paraíba		Jiangsu	Guainía	Kerala	México
Paraná*		Jiangxi	Guaviare	Madhya Pradesh	Michoacán
Pernambuco		Jilin	Huila*	Maharashtra*	Morelos
Piauí		Liaoning	La Guajira	Manipur	Nayarit
Rio de Janeiro*		Nei Mongol	Magdalena*	Meghalaya	Nuevo León
Rio Grande do Norte		Ningxia Hui	Meta*	Mizoram	Oaxaca
Rio Grande do Sul		Qinghai*	Nariño*	Nagaland	Puebla*
Rondônia		Shaanxi*	Norte de Santander*	NCT of Delhi	Querétaro
Roraima		Shandong	Putumayo*	Odisha	Quintana Roo
Santa Catarina*		Shanghai	Quindío	Puducherry	San Luis Potosí*
São Paulo*		Shanxi*	Risaralda	Punjab	Sinaloa
Sergipe		Sichuan*	San Andrés y Providencia	Rajasthan*	Sonora
Tocantins		Tianjin*	Santander*	Sikkim	Tabasco*
		Xinjiang Uygur*	Sucre*	Tamil Nadu	Tamaulipas*
		Xizang	Tolima*	Telangana	Tlaxcala
		Yunnan	Valle del Cauca	Tripura	Veracruz*
		Zhejiang	Vaupés	Uttar Pradesh	Yucatán
			Vichada*	Uttarakhand	Zacatecas
				West Bengal*	

Russia			United States		
Adygey	Khakass	Pskov	Alabama*	Nevada*	
Altay	Khanty-Mansiy*	Rostov	Alaska*	New Hampshire	
Amur	Kirov	Ryazan'	Arizona	New Jersey	
Arkhangel'sk*	Komi*	Sakha*	Arkansas*	New Mexico*	
Astrakhan'*	Kostroma	Sakhalin*	California*	New York*	
Bashkortostan*	Krasnodar	Samara*	Colorado*	North Carolina	
Belgorod	Krasnoyarsk*	Saratov	Connecticut	North Dakota*	
Bryansk	Kurgan	Smolensk	Delaware	Ohio*	
Buryat	Kursk	Stavropol'	District of Columbia	Oklahoma*	
Chechnya	Leningrad	Sverdlovsk*	Florida*	Oregon	
Chelyabinsk	Lipetsk	Tambov	Georgia	Pennsylvania*	
Chukot	Maga Buryatdan	Tatarstan*	Hawaii	Rhode Island	
Chuvash	Mariy-El	Tomsk*	Idaho	South Carolina	
City of St. Petersburg	Mordovia	Tula	Illinois*	South Dakota*	
Dagestan*	Moscow City	Tuva	Indiana*	Tennessee	
Gorno-Altay	Moskva	Tver'	Iowa	Texas*	
Ingush	Murmansk	Tyumen'*	Kansas*	Utah*	
Irkutsk*	Nenets	Udmurt*	Kentucky*	Vermont	
Ivanovo	Nizhegorod*	Ul'yanovsk	Louisiana*	Virginia	
Kabardin-Balkar	North Ossetia	Vladimir	Maine	Washington	
Kaliningrad	Novgorod*	Volgograd	Maryland	West Virginia*	
Kalmyk	Novosibirsk	Vologda	Massachusetts	Wisconsin	
Kaluga	Omsk*	Voronezh	Michigan*	Wyoming*	
Kamchatka	Orel	Yamal-Nenets	Minnesota		
Karachay-Cherkess	Orenburg*	Yaroslavl'	Mississippi*		
Karelia	Penza	Yevrey	Missouri		
Kemerovo*	Perm'*	Zabaykal'ye	Montana*		
Khabarovsk	Primor'ye		Nebraska*		

Note: *Oil-producing regions. To identify them as such, I use several national sources, as well as the reports from the U.S. Energy Information Administration. I excluded three regions of India due to lack of an economic activity indicator: Dadra and Nagar Haveli, Daman and Diu, and Lakshadweep.

A.2 Harmonization: Details

I first apply a cubic convolution resampling over the VIIRS data to then extract DMSP and VIIRS data averaging values using a spatial resolution of 10 km. I subtract 0.3 from each observation of VIIRS data (value identified as the threshold for low signal values by Li et al. 2017 and Wu and Wang 2019). Any negative value is set to zero after the subtraction.

Chapter 2 assumes the possible presence of two non-linear relationships in the overlapped 2013 year to obtain 2014–2019 DMSP annual composites using VIIRS data. This approach is less restrictive than choosing one type of association. The relationship is also allowed to vary at the country and regional level.

Power function specification:

$$x_{ntij}^{DMSP} = \hat{\alpha}_0 \left(x_{ntij}^{VIIRS}\right)^{\hat{\alpha}_1} + \hat{u}_{ntij} \tag{A.1}$$

Logarithm specification:

$$x_{ntij}^{DMSP} = \hat{\beta}_0 + \hat{\beta}_1 Log(x_{ntij}^{VIIRS} + 1) + \hat{v}_{ntij} \tag{A.2}$$

x^{DMSP} and x^{VIIRS} are the values of light intensity from DMSP, after the inter-annual calibration, and VIIRS sensors, respectively, in which n represents the number of pixels, t the year 2013, i the countries, and j the regions within the respective country. The estimated parameters would be $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\beta}_0$, and $\hat{\beta}_1$, while the errors are $\hat{u}$ and $\hat{v}$. I include unlit areas; therefore, I add one to x^{VIIRS} to avoid undefined values in the logarithm specification.

Both specifications are estimated using the non-linear least squares method by region. I use the least RMSE to select the coefficients to construct the simulated DMSP. Nevertheless, the DMSP values are limited to 63 DN from the source, while the fitted values might not. I solve this situation directly restricting the fitted values, such that:

$$\hat{x}_{ntij}^{DMSP} = \begin{cases} 63 & \text{if} \quad 63 < \hat{x}_{ntij}^{DMSP} \\ \hat{x}_{ntij}^{DMSP} & \text{otherwise} \end{cases} \tag{A.3}$$

A.3 Top-Coding Correction: Details

The procedure is adapted from Bluhm and Krause (2018). In this case, the general steps are as follows:

1. Identify the pixels with saturated lights (188,689 values $\geq 35DN$).
2. Produce a ranking of those pixels by country and year. Unlikely, I would have duplicates in the ranking considering the data extracted as averages, but in case of coincidences, a second ranking is generated based on the non-saturated values from VIIRS.
3. Generate corrected values assuming a truncated (bounded) Pareto distribution using the Bluhm and Krause (2018)'s shape parameter of 3/2 across all countries. I also ranked these values.
4. Replace the saturated lights with the corrected lights, matching the orders of the respective rankings.

Let the intensity of lights, x, be a random variable. The probability density function of x following a bounded Pareto distribution is:

$$f(x) = \frac{\alpha L^{\alpha} x^{-\alpha-1}}{1 - \left(\frac{L}{H}\right)^{\alpha}} \tag{A.4}$$

The distribution has three parameters: α determines the shape, L is the top-coding threshold, and H the upper bound. Chapter 2 assumes $\alpha = 3/2$, $L = 35$ and $H = 2000$. Applying the inverse-transform method and including the values of the respective parameters, with n representing the number of pixels, t the years in the sample and i the countries, I would have:

$$x_{nti} = (2000 \cdot 35)(u_{nti} 35^{\frac{3}{2}} + 2000^{\frac{3}{2}} - u_{nti} 2000^{\frac{3}{2}})^{-\frac{2}{3}} \tag{A.5}$$

Finally, corrected lights are generated by substituting into this expression values of u_{nti} equal to a random draw from a standard uniform distribution, that is, a $u \sim U(0,1)$ by country and year.

A.4 Summary Statistics

Table A.2: Summary Statistics (Pixel-Based), All Years

	Mean	Standard deviation	Minimum	Maximum	Observations
Regions: All					
Log(*Z*)	8.7465	3.4502	4.3464	14.9596	7,369
Log(*Sum of lights*)	7.8249	1.6327	-0.4931	11.2096	7,369
Log(*Intensity of lights*)	0.8274	1.8173	-11.0895	7.1085	7,369
Log(*Sum of lights/Population*)	-6.9469	1.2949	-11.0418	-1.9599	5,986
Regions: Non-oil					
Log(*Z*)	8.4708	3.4411	4.3464	14.9596	4,872
Log(*Sum of lights*)	7.5677	1.6293	-0.4931	10.8730	4,872
Log(*Intensity of lights*)	0.8114	2.0223	-11.0895	7.1085	4,872
Log(*Sum of lights/Population*)	-7.0867	1.2774	-11.0418	-1.9599	3,974
Regions: Oil					
Log(*Z*)	9.2844	3.4048	4.3870	14.8453	2,497
Log(*Sum of lights*)	8.3267	1.5190	3.3626	11.2096	2,497
Log(*Intensity of lights*)	0.8587	1.3291	-4.7495	4.2252	2,497
Log(*Sum of lights/Population*)	-6.6708	1.2852	-10.1829	-3.3831	2,012

Note: Z represents the economic activity indicators from Table 2.1