Planes, Trains, and Automobiles: What Drives Human-Made Light?

Dickinson, Jeffrey

2020

Planes, Trains, and Automobiles: What Drives Human-Made Light?

Jeffrey Dickinson[1,*]

[1]Graduate Institute of International and Development Studies

Abstract

This paper expands on our understanding of the lights-income relationship by linking the newest generation of nighttime satellite images derived from the Visible Infrared Imaging Radiometry Suite, VIIRS, to nationwide, panel data on population and income from 2012-2018 for both Brazil and the United States including 3,104 US counties, and 5,570 municípios. I leverage the quality and frequency of those data sources and the VIIRS lights images and find that nighttime light does indeed respond to changes in income. I find decreasing marginal effects of GDP on nighttime light as well as decreasing marginal effects of population on nighttime light, a result which holds across many specifications and that is robust to sub-sample analysis and placebo tests. Interactions among controls also appear to be present. Using sub-sample analysis, I also find that nighttime light does a poor job of capturing less-wealthy areas. Finally, I use a between-county estimator to identify the effects of time-invariant infrastructure features on night-time light. Roads, rail, ports, airports, and border crossings I find to be strong contributors to increases in light.

JEL Codes O51, C82, R10, R11, R12

Keywords: night-time light, GDP, population, infrastructure, regional development, Brazil, USA

*Correspondence: jeffrey.dickinson@graduateinstitute.ch, Department of International Economics, Graduate Institute of International and Development Studies, CP 1672 - CH-1211 Genève 1, Suisse. I express my heartfelt thanks to my two doctoral supervisors, Jean-Louis Arcand, who introduced me to the subject of nighttime lights and impact evaluation, and Nicolas Berman, who was my first graduate econometrics teacher. I am grateful for their extensive support, academic and otherwise, and many, many thoughtful comments. All remaining errors are my own.

1 Introduction

Using the newer Suomi-NPP satellite, Nasa collects high-resolution imagery of the earth at night. The newer images, captured on the Visible Infrared Imaging Radiometry Suite (VIIRS), offer major advancements over the previous generation of nighttime images that originated from defense department weather satellites. The literature using nighttime lights satellite images as a proxy measure for human activity dates back to the 1970's but the watershed papers relating nighttime light to economic variables were those by Henderson et al. (2012), and Chen and Nordhaus (2011). These two papers proposed that nighttime lights could be used as a proxy indicator for income, and they analyzed the correspondence between national accounts data and night-time lights at the highest level of aggregation, the country, finding a fairly strong relationship between income and lights. The authors in Henderson et al. (2012) faced sharp limitations with their data: the reference national accounts data from many low-income countries could be noisy making identification of the exact parameters linking income, GDP, and population difficult and, worse, potentially causing omitted variable bias. Furthermore, the data from the previous generation of satellites were top-coded, and unable to record light values beyond a certain integer, 63. This translated into many dense and bright areas being top-coded implying loss of information. The new images no-longer face this limitation as the new radiometry suite has been custom-built to capture nighttime imagery. Recent work, using high-quality cross-sectional data from Sweden, has suggested that light growth is closely linked with population movements more than with fluctuations in income (Mellander et al., 2015). Levin and Zhang (2017) also utilizes data from the newer VIIRS satellite, the same lights dataset used in this paper, and analyzes lights-income relationship for all the urban areas on the globe (n=4,153) in the months of January 2014 and July 2014. They find that lights are more closely related with *national* income per capita than with population.

With respect to papers whose analysis utilizes nighttime lights at a more detailed level, e.g. at a higher spatial resolution, the literature is been growing. Hodler and Raschky (2014) examine the presence of stronger growth in regions associated with the leader of a country, and find a significant result. Mellander et al. (2015), perhaps the paper most similar in spirit to this one, is a well-cited paper which examines the relationship between economic activity, population, enterprise density, and nighttime light in Sweden using cross-sectional analysis. The authors find that light growth corresponds most to nighttime population density (population), rather than daytime enterprise density. Mellander et al. (2015) also argue that night-time light is only weakly correlated with income, although in their OLS regressions night-time light appears to increase by 0.424 units with an increase of one unit of Total Wage Incomes. Two new papers have recently been published using night-time lights for localized analysis. One measures the effects on light of flooding in cities around the globe, and finds that low-lying areas in cities recover as fast as other areas, and there appear to be no permanent effects of flooding on city development (Kocornik-Mina et al., 2020). Frick et al. (2019) uses night-time lights data to analyze the effect of special economic zones on economic activity. They find that key determinants to the success of special economic zones was linked with pre-existing industrial infrastructure in the surrounding area, and the presence of large markets in which to sell outputs. Bleakley and Lin (2012) uses night-time lights from the years 1996-7 to test for path-dependence around certain natural water features in the United States. The authors find that portage sites, sites where, in the past, transport boats could not pass and thus cities arose, are likely to still be of a substantial size around 100 years after the portage sites were relevant. Smith and Wills (2018) is a recent paper which leverages the global nighttime lights coverage to estimate the fraction of the population below the poverty line, and they find that spillovers from economic activity rarely disseminate to rural populations. An overview of the capabilities and some applications of night-time lights data can be found in Donaldson and Storeygard (2016).

The United States, with approximately 3,104 counties, in contrast with the data from Sweden used in Mellander et al. (2015), is a much larger landmass and total population (10m vs. 350 m), and has substantial heterogeneity with respect to landmass and shape, demographic composition, population density, and geographical characteristics such as mountains, lakes, rivers, and coastlines. This is evident when we consider places like California, which has only 58 counties per 40m citizens, Alaska, which is has enormous counties but is sparsely populated, Arizona, which is mostly desert and borders Mexico, Washington which has dense deciduous and evergreen forest, mountains, and a shared border with Canada, as well as Hawaii, an island halfway between the US and Japan in the Pacific ocean.

Brazil, in contrast, is a country with 211 million people,[1] and at the second administrative boundary level, has 5,570 municípios. The name translates to 'municipalities,' and they are, on average, smaller than counties, though there is overlap between município size and county size. There is also substantial heterogeneity in Brazilian municípios ranging from the unique coastal city of Rio de Janeiro to Manaus, in the middle of the Amazon. Brazil has dense and poor areas to a much larger extent than the USA. Since the two countries combined cover many heterogenous county and minicípio types, analyzing these two samples combined as well as separate I believe is a highly informative exercise. Combining the USA and Brazil samples allows me to leverage more than 55,000 observations, 21,728 from the USA and 33,414 from Brazil, and results with the two samples combined are shown alongside results from the separate samples throughout the paper.

The principal contributions of this paper are: to further understanding of the lights-income-population nexus by linking lights to administrative panel data of high quality, which are available at a fine spatial resolution. Another contribution is to demonstrate that the lights-population-GDP variables may be endogenous, which may lead to unreliable estimates in which case they should be utilized with caution. Another contribution is clarifying the existence of pronounced non-linear relationships appear in the aggregate and restricted-sample relationships including interaction terms among control variables. This also indicates that estimations which omit those terms may be omitting important variables. Another important contribution of the paper is to estimate the effects of time-invariant infrastructure features on light.

Utilizing the full size (n=55,142) of the dataset I am able to conduct extensive sub-sample

[1]https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?=t=resultados

analysis, as well as estimate the effects of time-invariant infrastructure features on nighttime light. I find that nighttime lights tends to be correlated more strongly with income in wealthier, larger counties, and the direct effect of GDP on nighttime light is often unreliably estimated indicating endogeneity. I also compare the nighttime light measure alongside electrical consumption data at the county level in California over the sample years. Previous authors have suggested that electrical consumption data may be of a similar value to NTL as a proxy indicator (Mellander et al., 2015; Henderson et al., 2012). I find that electrical consumption does correlate with higher levels of GDP and population, though in the within-county model we only see an effect of increases in the population on an increase in non-residential light.

The rest of the paper will proceed as follows: section 2 motivates the methodology used in the paper. Section 3 discusses the data sources and availability including a detailed description of the VIIRS nighttime lights data. Section 4 presents the results, and section 5 concludes.

2 Methodology

The main approach of this paper is to use panel-data tools reveal the links between population growth, income growth, and night-time light as measured. Using night-time light as the dependent variable makes the most sense, I argue, in the context because the satellite images from the VIIRS are a little noisy, while they are very precise in the dimension of how they record the texture of activity across space.[2] The general model, a night-time light production function, states simply that night-time light, as measured from the VIIRS sensors is a function of income, population, and other factors:

$$NTL_{ct} = \beta_1[GDP_{ct}] + \beta_2[POP_{ct}] + \beta_3[\text{Area}_{ct}] + \alpha_c + \phi_{st} + \varepsilon_{ct} \tag{1}$$

Where c indexes the county or município, t indexes the year, and α_c are the county/município fixed effects. The area variable controls for any potential relationship between the size of the county and the measurement of the lights that may not be captured by the income and population variables. Based on previous papers such as Hu and Yao (2019), there is reason to believe that income and population may not enter the night-time light production function linearly. This is an important consideration for our purposes as nonlinearities may mask desired effects of interest. In that case I will also estimate a translog specification, which includes squared terms and interaction terms among all three key independent variables. The intuition behind the squared terms is that there could be strongly diminishing effects in the way that income and population enter the production function. The interaction terms are included to capture the possibility that the lights-income or lights-population relationship could be stronger in larger counties or smaller ones. The third main variable besides income and population being the area of the county, which controls just for the total size of the county, as there is quite a

[2]This will be discussed further in the data section. The night-time lights images must undergo processing in order to remove image distortions which are orthogonal to changes in human-made light.

large variation. The second potential specification is therefore the following:

$$NTL_{ct} = \beta_1 X + \beta_2(X^2) + \beta_3(x_1 \times x_2 ...) + \alpha_c + \psi_{st} + \varepsilon_{ct} \tag{2}$$

The first term is the normal, log-transformed variable, the second term is the squared version of all control variables, and the third term is the interaction of all control variables.

Between-county Estimation

There are certain geographic characteristics of the counties and municípios which we would like to analyze, but it is difficult because infrastructure features are largely invariant within the sample period, 2012-2018. In order to obtain identification of time-invariant features, all variables are collapsed to their group means. This procedure is similar to the strategy employed in Henderson et al. (2012), who also employ the within-transformed country-level data, and then in their case they used long-differences instead of group means. Identification of the effect of the infrastructure or geographic features then comes from comparing counties which have infrastructure or features exclusively to other counties within the same state-year that lack those features. Given the size of the sample (n=55,142 county-years and município-years) and the survey period I feel this is the most appropriate approach to consider the effects of geographic variables. The between estimator can be represented in the following form:

$$\bar{y}_i = \alpha + \beta \bar{x}_i + \phi \bar{x}_i^2 + \gamma \bar{x}_{ij} * \bar{x}_{ik} + \theta G_i + \bar{\epsilon}_i \tag{3}$$

where $\bar{y}_i$ is the mean value of nighttime light in county or município i, $\bar{x}_i$ is the mean value of the control variable for the county or município, G_i is an indicator for geographic features, and the main parameters of interest are then β, θ, and γ.

3 Data

Table 1 details years of data availability. The LandScan data has the best coverage through time, while the VIIRS nighttime lights series starts only in 2012. The binding constraint on our sample is therefore the population data as we have no estimates for population at the county level past 2018, and I am able to leverage the years 2012-2018.

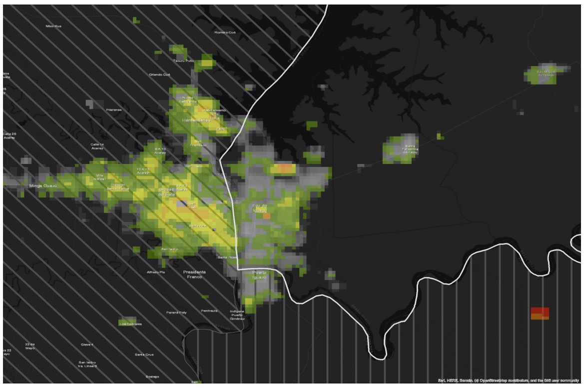
		Source	Years Available
GDP	USA	BLS	2001-2018
	Brazil	IBGE	2002-2017
Population	USA	ACS/census	2009-2018
	Brazil	IBGE	1975-2017
Lights	Both	NoAA/NASA	2012-present
Landscan	Both	ORNL	2012-2018

Table 1: Data Availability

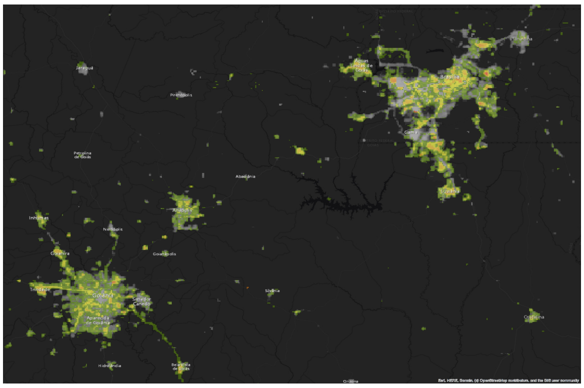
3.1 BLS/IBGE GDP Data

Over the past few years the Bureau of Labor Statistics (BLS) has been releasing local-area calculations for gross domestic product. In the BLS GDP statistics, county-level GDP is calculated using the income approach. Based on the availability of data, the Bureau of Economic Analysis (BEA) utilizes the income method for calculating county-level GDP. *"GDP is computed as the sum of compensation of employees, taxes on production and imports less subsidies, and gross operating surplus. The initial regional estimates are then scaled to the national estimates so that all BEA estimates are reconciled"* (Aysheshim et al., 2020). Principal sources of the county-level GDP data are the Department of Labor's Quarterly Census of Earnings and Wages, aircarrier traffic statistics, DOT surface transportation data, bank branch deposits, and other proprietary government sources. A full accounting of all sources and information used in the calculation of GDP at the county level can be found in Aysheshim et al. (2020). There is substantial between-county variation in the GDP data: some counties produce millions of dollars, while others produce well under 100k per annum.

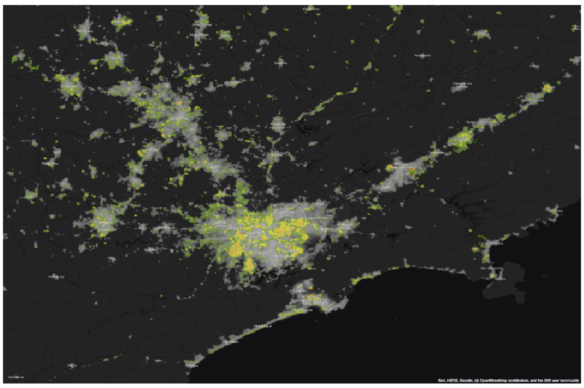
On the Brazilian side the Brazilian GDP data comes from the Instituto Brasileiro de Geografía e Statística (IBGE) and the data are compiled from governmental and other administrative data sources, very similar to the USA GDP estimates.[3]



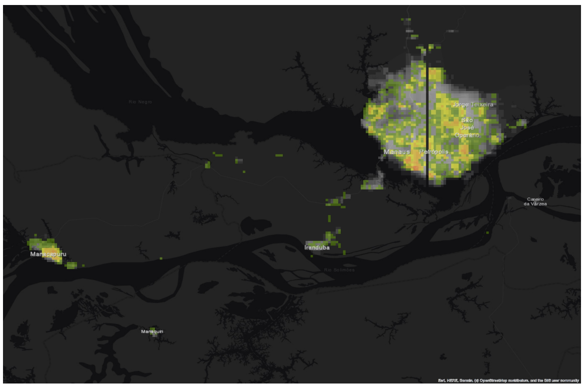
(a) Foz do Iguaçu, PR



(b) Brasilia, DF



(c) Sao Paolo, SP



(d) Manaus, AM

Figure 1: Night-time Lights of Four Major Brazilian Cities;
Layers: Basemap: Open Street Map, CC License; Night-time Lights Annual Image (2019); Changes in NTL 2012-2017 - Green = small change, Red = large change

[3]The full details of all sources and methods for the production of the Brazilian GDP estimates can be found on the IBGE website

3.2 ACS/IBGE County-Level and Município-level Population Data

Population estimates come from American Community Survey (ACS) 5-year estimates of the county-level population. These are calculated using data sampled from the county on a rolling basis over the course of 5 years. ACS data are the main survey data that are collected from communities in the United States in the intercensal period.

Like the GDP estimates, the Brazilian population estimates also come from the IBGE, and they are based on the Brazilian population census which took place in 2000 and 2010, adjusted for changes in between.

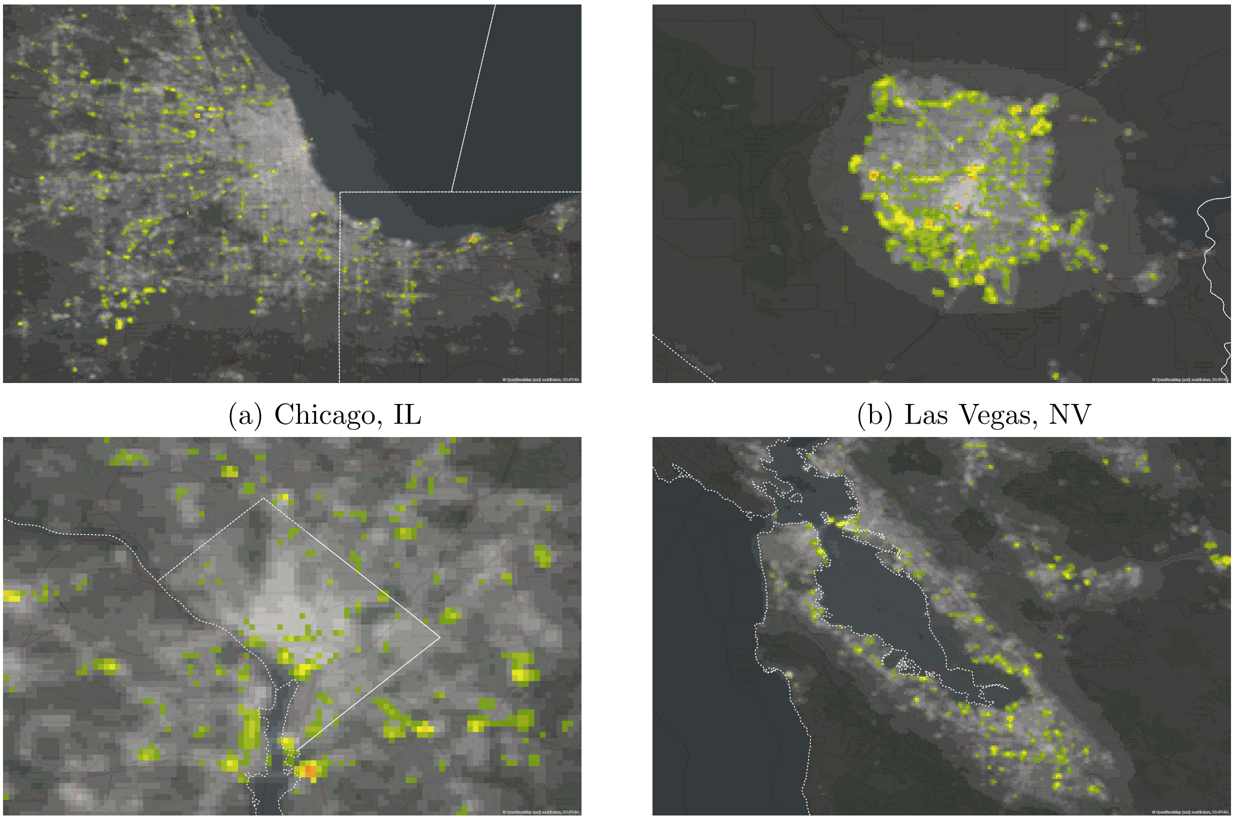
3.3 LandScan Gridded Population Data

LandScan gridded population data is a global population dataset in the form of an integer-based raster, with annual rasters available from 2001-2018. The population is inferred using an algorithm and a mix of sources, with one principal source being high-resolution daytime satellite imagery of human settlements. The LandScan dataset is popular, and has been used in other economics research when comparable administrative population data are not available.

3.4 VIIRS Night-time Lights Data

The Suomi-NPP Satellite project, which started in 2011, is a joint civilian venture of the United States National Aeronatuic and Space Administration (NASA), the Department of Defense, and the National Oceanographic and Atmospheric Administration. The Visible Infrared Imaging Radiometer Suite (VIIRS) is intended to capture human-made light and overcomes many limitations of the previous Defense Meteorological Satellite Program (DMSP) satellite images. The newer Suomi NPP satellite, which contains the VIIRS, has an automatic gain sensor which adjusts to allow greater sensitivity, meaning the device can better capture much lower and higher levels of light (Elvidge et al., 2017). The resolution of the new VIIRS images, available from 2012-2020, with data available on a daily frequency or in monthly composite forms, is extremely high, with pixels being around 742m across compared to the DMSP pixels which were 3km across (Carlowicz, 2012; Elvidge et al., 2017). This sensitivity is of extreme interest to researchers in attempting to pinpoint precise locations which are centers of economic activity, and will reduce limitations around night-time lights data coming from heavily saturated urban areas. The Suomi-NPP satellite flies over the earth around 1:30am and 1:30pm local time each day and captures images using the spectroradiometer, a device similar to the capture device in a digital camera (Carlowicz, 2012). Raw data from the sensor are then processed to remove non-human generated disturbances such as aurora borealis, stray light, natural fires and other light which could potentially introduce noise. A detailed accounting of the processing of the data can be found in Elvidge et al. (2017).

Some examples of night-time lights images of major Brazilian cities are shown in figure 1, and US cities are shown in Figure 2. Long-run changes in night-time light are shown in green-red colors to demonstrate intensity. First in the top left image of figure 1 we can see the



(a) Chicago, IL

(b) Las Vegas, NV

(c) Washington, DC

(d) San Francisco, CA

Figure 2: Night-time Lights of Four Major US Cities;
Layers: Basemap: Open Street Map, CC License; Night-time Lights Annual Image (2019); Changes in NTL 2012-2019 - Green = small change, Red = large change

city of Foz do Iguacu, PR Brazil, which straddles the border with Paraguay, on the left, and Argentina, to the south, at the site of an important hydroelectric dam, the Itaipu dam, on the Brazil-Argentina-Paraguay border; development on the Paraguayan side appears to be more aggressive over the 2012-2017 period. We see much more development on the Paraguayan side than on the Brazilian side. Changes in both the extensive and intensive margins are visible on the Paraguayan side, while on the Brazilian side there is much less change at the extensive margin and light/growth appears to be condensed along the highway. In the top right corner of the figure, panel b shows Brasilia, DF which has experienced a relatively rapid period of development relative to other parts of Brazil, in the top right hand corner of panel b, stretching down to Goîana in the bottom left corner with Anápolis visible in between. The bottom left corner is Sao Paolo, SP, by far the most populated region of Brasil with 48.6m persons, which appears to have substantial development and sprawl along the coastline and the highway corridor. Last in panel d we have Manaus which is a Brasilian city in the rainforest. The increases in the intensive margin, light intensity, are clearly much more intense than changes in the extensive margins, which would be indicated by outward expansion of nighttime light. For the american cities in figure 2, Chicago, IL is shown in the upper left panel, panel a, and is seen to be quite spread out over space. Las Vegas, NV, in panel b, is an interesting example because of its intensity relative to the darkness of the nearby unpopulated desert. Panel c shows how Washington, DC illustrates that, despite high density of lights, changes in light intensity can

VARIABLES	(1) Total NTL	(2) Total NTL	(3) Total NTL	(4) Total NTL
Commerical Elec. Cons.	0.712***			
	(0.0178)			
Residential Elect. Cons.		0.772***		
		(0.0243)		
Combined Elect. Cons.			0.763***	0.593
			(0.0183)	(0.557)
Observations	406	406	406	406
R-squared	0.869	0.806	0.868	
Number of Counties				58

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: California Nighttime Lights (log) Regressed on the Log of Electrical Consumption

still be distinguished at a high resolution. The dark red spot just south of Washington, DC is National Harbor, an area of major development for the DC metropolitan area over the last few years. The major development inside DC over that period was the Southwest Waterfront, which can also be seen as the glowing yellow dot at the southern tip of DC where the Potomac River meets the Anacostia. Lastly, one of the wealthiest, most expensive, and most productive regions in the country is depicted in Northern California from Berkeley to San Jose, revealing pockets of development along the way. Tables 18 and 19 show the counties with the most and least light, and are included in the appendix. The variance in light is substantial, from Robertson County, KY, the county with the least total light, to Yukon-Koyukuk County, AK with the most light.

3.5 California Electrical Consumption Data

California's state energy agency, California Energy Commission, makes available electrical consumption data at the county level for all counties in California.[4] These data are available at the county level from 1990-2018. They are administrative in nature and are therefore, to the best of my knowledge, do not represent a sample of electrical consumption data. A regression of NTL on electrical consumption can be seen in table 2. As we can see, nighttime light is strongly correlated with electrical consumption, slightly more so with non-residential electrical consumption.

3.6 Infrastructure Data

Infrastructure data, including the location of ports, rail, navigable waterways, and Fortune-500 business headquarters have been collected from the U.S. federal government's Homeland Infrastructure Foundation Level Database (HIFLD) website, which is funded under the Department

[4]https://ecdms.energy.ca.gov/elecbycounty.aspx

of Homeland Security. Airport locations were taken from open data sources.[5] Data on primary roads, which includes interstates and principal highways, was collected from the US Census Department.

4 Results

[5]https://ourairports.com/

VARIABLES	N	mean	median	sd	min	max
Total Nighttime Light	406	54822	17507	112144	755.6	822111
BLS GDP	406	41730000	7615000	97600000	47224	710900000
LS Population	406	668138	181767	1453000	1140	10140000
ACS Population	406	669915	181536	1452000	1057	10120000
miles^2	406	2727	1554	3097	48.56	20118
km^2	406	7063	4024	8020	125.8	52104
Non-residential Elec. Cons.	406	3315	781.4	7021	4.008	49193
Residential Elec. Cons.	406	1585	553.2	3090	9.291	21162
Total Elec. Con.	406	4901	1474	10032	13.89	69946

Table 3: Summary Statistics of Variables Used in Electrical Consumption Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Total NTL	Total NTL	Total Elec	Total Elec.	Resid. Elec.	Resid. Elec.	Comm. Elec.	Comm. Elec.
Area	0.486***		0.147***		0.209***		0.0472***	
	(0.0206)		(0.0143)		(0.0205)		(0.0133)	
BLS GDP	0.551***	0.261***	0.235***	0.0419	0.392***	0.0993	-0.00390	-0.00551
	(0.0572)	(0.0790)	(0.0272)	(0.0337)	(0.0503)	(0.131)	(0.0484)	(0.0382)
ACS Population	0.0974	-1.239	0.672***	0.525*	0.555***	0.374	0.878***	0.712***
	(0.0637)	(0.926)	(0.0292)	(0.300)	(0.0562)	(0.393)	(0.0545)	(0.178)
Constant	-3.670***		-5.638***		-7.688***		-4.616***	
	(0.296)		(0.182)		(0.274)		(0.213)	
Observations	406	406	406	406	406	406	406	406
R-squared	0.922		0.981		0.956		0.964	
Number of Counties		58		58		58		58
County FE		yes		yes		yes		yes

Columns 1,3,5,7: clustered standard errors (county) in parentheses
Columns 2,4,6,8: cluster-robust standard errors (county) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: California Electrical Consumption Regressions

4.1 California Electrical Consumption Regressions

Table 3 contains the summary statistics of variables used in the electrical consumption regressions, and table 4 shows the results of regressions those regressions. The availability and granularity of the California data permit the direct comparison of the value-added of night-time lights over electrical consumption data. Columns 1-2 are the regression of only the California night-time lights using the same set of parsimonious controls as earlier. We see in column 1 and 2 that nighttime lights tracks with BLS GDP in California as well as the area, and this relationship is significant both in the global and the within regressions. With respect to the electrical consumption data, they track more closely with increases in the population as we see in column 3, and in column 4, which is the within-county transformed regression, none of the independent variables are significant. Looking at columns 5-8 which are residential (5-6) and non-residential (7-8) electrical consumption separated out, we see that population moves with electrical consumption, but that income moves with electrical consumption less, and income is only statistically significant in column 5, global-OLS with year fixed effects.

4.2 Cobb-Douglas Estimates

Summary statistics for the principal regression variables can be found in table 19 in the appendix. The county population variable, LandScan version, the smallest county has 85 residents, Loving, TX while the largest has 10,140,000, Los Angeles, CA. The ACS 5-year estimates are very similar. The results presented in Table 5 are the estimates of the Cobb-Douglas nighttime light production function.

All variables are in log form and all columns include state·year fixed effects. The first two columns in table 5 represent the estimates using the combined datasets, column 1 corresponds to estimates using administrative population data, column 2 utilizes LandScan satellite-inferred population data. The next two columns are the estimates with the sample restricted only to the US, and the final two columns are the same model with the sample restricted exclusively to the Brazilian data. All columns contain both county/município fixed-effects, and state-year fixed effects. It is clear from the first two columns that all of the control variables except for area are significant. We see in column 1 and 2 that, according to this specification we see that light is correlated with GDP, though in this context it appears light responds more strongly to population changes than it does to light. Looking at the US columns, it is a similar story, though we see a slightly stronger overall effect of population on nighttime light. Turning to Brazil the effects of GDP on light is slightly stronger. In terms of the effect of population on light, the effect of population on light for Brazil is not significant at traditional levels, but it is estimated to be much lower at .015 meaning an increase of population of 1% results in a .015% increase in light.

VARIABLES	(1) NTL	(2) NTL	(3) NTL	(4) NTL	(5) NTL	(6) NTL
Area	0.232**	0.227**			0.226**	0.226**
	(0.114)	(0.114)			(0.115)	(0.114)
GDP	0.0453***	0.0453***	0.0285	0.0277	0.0907***	0.0906***
	(0.0151)	(0.0151)	(0.0178)	(0.0171)	(0.0158)	(0.0158)
Pop	-0.0126		-0.463***		0.0191	
	(0.0451)		(0.122)		(0.0456)	
LS Pop		0.0111		-0.363***		0.0156*
		(0.00797)		(0.0611)		(0.00804)
Observations	55,142	55,142	21,728	21,728	33,414	33,414
County FE	Yes	Yes	Yes	Yes	Yes	Yes
State*year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Admin Areas	8,674	8,674	3,104	3,104	5,570	5,570

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Cobb-Douglas Light Production Function

4.3 Translog Estimates

The next table, table 6, represents the translog functional form. The arrangement is the same as the previous regression table with columns 1-2 corresponding to the combined estimates, followed by 3-4 being the USA estimates, and 5-6 being the BRA estimates. This specification includes the second-order terms as well as interactions among all independent variables. Now it is evident that an increase in GDP corresponds to an increase in light, with the effect varying in magnitude but statistically significant across all four columns. Nighttime light is also strongly increasing in the overall population. The second-order term for GDP is significant, and small in magnitude, though only significant at the standard levels in the case of Brazil. The estimated second-order effects for the combined USA data are very small in magnitude. The same term for population is larger, and appears to be well-estimated across columns, though in the USA the administrative data yields a much larger estimate than the LandScan data in column 4. The area×control interaction terms are statistically significant in the case of the population×area variable it is negative, whereas GDP×area is significant in one case and positive. This means, in other words, that the larger the area of the county, the smaller the magnitude of the relationship between population and light, while the inverse is true for GDP. Lastly, the population×GDP variable is significant, though the effect is small, the effect is positive in the USA and estimated to be negative in Brazil. A positive GDP×population variable is interesting because the size and significance of this estimate indicate that with a higher population, the relationship between GDP and light is stronger, and in counties with a higher GDP, there is a stronger relationship between population and nighttime light.

VARIABLES	(1) NTL	(2) NTL	(3) NTL	(4) NTL	(5) NTL	(6) NTL
Area	0.497	0.236			0.653	0.149
	(0.590)	(0.580)			(0.589)	(0.578)
GDP	0.315***	0.357***	0.105	0.275	0.546***	0.581***
	(0.0942)	(0.0917)	(0.135)	(0.207)	(0.125)	(0.113)
Pop	1.399***		1.441*		0.254	
	(0.249)		(0.772)		(0.343)	
GDP^2	-0.000336	-0.000147	0.000458	0.00142	-0.0383***	-0.0216***
	(0.00201)	(0.00186)	(0.00256)	(0.00277)	(0.0127)	(0.00502)
Pop^2	-0.0559***		-0.149***		-0.0333**	
	(0.0120)		(0.0291)		(0.0160)	
$Area^2$	0.0190	0.00736			0.0131	0.00229
	(0.0451)	(0.0445)			(0.0443)	(0.0438)
Area*Pop	-0.0400***		0.188**		-0.0506***	
	(0.0144)		(0.0825)		(0.0142)	
Area*GDP	-0.00992	-0.0173**	0.0145	-0.00776	-0.00976	-0.00734
	(0.00779)	(0.00702)	(0.0121)	(0.0138)	(0.0118)	(0.0109)
Pop*GDP	-0.0201***		-0.0228***		0.0597*	
	(0.00557)		(0.00689)		(0.0306)	
LS Pop		0.229***		1.191**		-0.0361
		(0.0685)		(0.478)		(0.0775)
LS Pop^2		-0.00294		-0.0364**		-0.0104**
		(0.00425)		(0.0156)		(0.00511)
Area*LS Pop		0.0106		-0.0714		0.0147*
		(0.00815)		(0.0504)		(0.00837)
LS Pop*GDP		-0.0206***		-0.0245***		0.0107
		(0.00544)		(0.00880)		(0.00856)
Observations	55,142	55,142	21,728	21,728	33,414	33,414
County/município FE	Yes	Yes	Yes	Yes	Yes	Yes
State×year FE	Yes	Yes	Yes	Yes	Yes	Yes
# of Admin Areas	8,674	8,674	3,104	3,104	5,570	5,570

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Translog Light Production Function

Tercile	USA Counties	BRA Municípios
	Area	
1	511	17,878
2	8,771	9,615
3	12,446	5,934
Total	21,728	33,427
	Pop	
1	4,366	14,016
2	6,643	11,738
3	10,719	7,661
Total	21,728	33,415
	GDP	
1	343	18,027
2	6,521	11,849
3	14,831	3,539
Total	21,695	33,415

Table 7: Counties and Municípios by Terciles

4.4 Terciles of Area

The following tables 8-10 will take a similar format where the first three columns show estimates using the combined USA and Brazil samples. The middle three columns of table 8 represent the estimates with the sample restricted exclusively to the USA, and the last three columns represent the same estimates with the model applied to the Brazilian sample. The first in the series breaks down the administrative districts into terciles based on the size of the administrative district (counties and municípios). Column 1 is the smallest tercile of counties/municípios, which corresponds to counties and municípios less than 467 square km, the middle tercile corresponds to counties and municípios greater than or equal to 467 square km and less than 1495 square km, and the largest tercile is those municípios and counties greater than 1495 square km.

Looking across the row for GDP we see very few estimates are significant, and the effect size varies widely and the instability of the parameter estimates may be the result of endogeneity among light, GDP, and population. With respect to population, the next row down, except for columns 2, 5, and 6, the effects are large and positive and mostly statistically significant. Turning to the second-order terms, GDP^2, which is unlikely to be endogenous to other variables and light, is much better estimated with the effects being negative and fairly small. With respect to the population^2 terms, the estimates are predominantly negative, and the statistically significant estimates range from -.04 to -.172.

	Combined			USA			BRA		
Terciles of Area	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Dep. Variable	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	0.356	-3.734	5.135*				0.330	-4.206	5.261*
	(1.204)	(5.022)	(2.951)				(1.211)	(5.048)	(2.943)
GDP	0.166	-7.55e-05	0.0982	1.131*	0.106	0.0820	0.164	0.0713	0.727
	(0.155)	(0.311)	(0.131)	(0.612)	(0.331)	(0.149)	(0.156)	(0.432)	(0.442)
Pop	3.075***	-2.333	1.301***	1.079	5.603***	0.471	3.034***	-4.744**	-0.891
	(1.032)	(1.617)	(0.200)	(2.289)	(1.994)	(1.276)	(1.057)	(2.213)	(0.869)
GDP^2	-0.0112	-0.0584***	0.00120	-0.00827	-0.0656***	0.000215	-0.0113	-0.0652***	-0.0912**
	(0.0118)	(0.00905)	(0.00262)	(0.0251)	(0.0110)	(0.00293)	(0.0119)	(0.0123)	(0.0463)
Pop^2	-0.172***	-0.0629	-0.0476***	-0.139*	-0.105***	-0.175***	-0.171***	-0.0277	0.00565
	(0.0499)	(0.0457)	(0.0163)	(0.0807)	(0.0326)	(0.0353)	(0.0518)	(0.0871)	(0.0269)
$Area^2$	-0.0398	0.0565	-0.280				-0.0410	-0.0216	-0.266
	(0.0801)	(0.361)	(0.174)				(0.0799)	(0.363)	(0.175)
Area×Pop	0.0207	0.255	-0.0344***	0.349	-0.700**	0.359***	0.0252	0.461*	-0.0974***
	(0.122)	(0.208)	(0.0115)	(0.302)	(0.289)	(0.129)	(0.125)	(0.251)	(0.0198)
Area×GDP	0.00627	0.0705*	0.0172	-0.0670	0.131***	0.0159	0.00581	0.0291	0.0292
	(0.0225)	(0.0396)	(0.0116)	(0.0706)	(0.0505)	(0.0131)	(0.0228)	(0.0546)	(0.0457)
Pop×GDP	0.0155	0.116***	-0.0252***	-0.0480	0.0783***	-0.0218**	0.0160	0.156***	0.153*
	(0.0315)	(0.0232)	(0.00779)	(0.0538)	(0.0241)	(0.00852)	(0.0316)	(0.0377)	(0.0890)
Observations	18,388	18,386	18,368	511	8,771	12,446	17,877	9,615	5,922
County/Mun. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Admin areas	3,066	2,871	2,775	73	1,253	1,778	2,993	1,618	997

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Estimates by Tercile of Area

4.5 Terciles of GDP

Looking at the next table, table 9, which follows the same format but now the terciles are terciles of GDP rather than administrative area. Tercile 1 corresponds to the poorest households, tercile 3 corresponds to the wealthiest. Looking across the column for GDP we see that the wealthiest counties are the ones which appear to be strongly linked with growth in lights. In the case of the USA, all terciles of income are strongly linked with light and they are statistically significant at the highest levels. The effect of larger population on light is estimated to be negative for the poor and middle-class tercile, then positive for both countries and the combined estimates in the wealthiest tercile. This is intriguing because we have strong evidence of an inverse relation for low-levels of income and positive one for high-levels. This relationship will be explored further in the subsequent tables where the individual terciles of income are broken down further. The second-order terms are very similar to the previous table. Again, this stability of the parameter estimates of the second-order terms is likely driven by the fact that these parameters are not endogenous to population or GDP. The estimates remain low, and are fairly tightly with estimates of beta ranging from .040-.065. The population-squared term is positive in the USA sample for the lowest tercile, though for the wealthiest tercile light is decreasing strongly in population, with estimates for beta and appears to be quite tightly estimated ranging from -.167 to -.230. With respect to the area$\times$var controls, I would like to note that in the case of the USA data, there is an effect there that appears to be well-estimated and the estimates range between .046-.061. Last, the final interaction term which represents the interaction of population and GDP, is in many cases significant, small, and positive meaning a higher population increases the strength of the GDP-lights relationship and a higher GDP increases the strength of the Population-lights relationship.

4.6 Terciles of Population

Next up we have table 10 which breaks down the sample into terciles of population. The smallest tercile is counties and municipios less than 10k persons, the middle is 10k to 25k, and the largest is counties and municípios above 25k. Looking across the second row, which corresponds to the effect of GDP on nighttime light, we can see the estimates are consistently positive, though only significant for the first two terciles in the combined, and the second and third tercile in the Brazilian sample. With respect to the population effect estimates, the combined estimates show a positive effect across all terciles, while the USA sample reveals something slightly different with the middle-population tercile having a negative effect on nighttime light, and the top tercile the effect of population on light is negative. The GDP^2 term is estimated to be smaller and negative as the previous tables. As mentioned before the estimates are fairly stable across population terciles. The population^2 term reveals a negative relationship for the top tercile of population, counties and municípios ¿ 25,000 persons. For the middle tercile, the second-order term for population is positive, and although not statistically significant at standard levels, the pattern fits with previous tables in the sub-sample analysis. Lastly, of interest is the population*GDP interaction term which is estimated to be negative for the US and positive for

Brazil, though the estimates are not all significant. It is of note that in the Brazilian estimates the effect size is estimated to be exactly the same for the bottom and top tercile of population at 0.0852.

	Combined			USA			BRA		
Terciles of GDP	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	0.361	0.987	1.324				0.181	1.174	1.232
	(1.004)	(0.941)	(1.159)				(1.014)	(0.959)	(1.401)
GDP	0.296	0.434	1.122***	2.624	0.906**	1.120***	0.220	0.379	0.833***
	(0.418)	(0.298)	(0.203)	(2.482)	(0.372)	(0.277)	(0.433)	(0.376)	(0.296)
Pop	-1.792	0.142	3.181***	-12.22**	-5.044**	2.361***	-0.608	0.00642	4.166***
	(1.746)	(0.273)	(0.697)	(5.752)	(2.251)	(0.889)	(2.144)	(0.325)	(1.522)
GDP^2	-0.0447*	-0.0361**	-0.0472***	-0.0626	-0.0654***	-0.0535***	-0.0465*	-0.0399**	-0.0545***
	(0.0261)	(0.0151)	(0.0114)	(0.0910)	(0.0192)	(0.0144)	(0.0263)	(0.0186)	(0.0137)
Pop^2	0.00256	-0.0335*	-0.175***	0.795***	0.120	-0.167***	-0.0856	-0.0368	-0.230***
	(0.108)	(0.0180)	(0.0343)	(0.245)	(0.0894)	(0.0397)	(0.141)	(0.0236)	(0.0753)
Area^2	-0.0496	0.00401	-0.121*				-0.0528	0.00238	-0.121*
	(0.0548)	(0.0693)	(0.0673)				(0.0546)	(0.0692)	(0.0697)
Area×Pop	0.0846	-0.0482***	0.0341	0.479	0.345***	0.129	0.115	-0.0563***	0.0429
	(0.106)	(0.0121)	(0.0592)	(0.601)	(0.125)	(0.0874)	(0.108)	(0.0128)	(0.0952)
Area×GDP	-0.0346*	-0.0111	0.0144	0.0135	0.0614***	0.0457*	-0.0389*	-0.0179	0.0109
	(0.0189)	(0.0134)	(0.0132)	(0.0565)	(0.0189)	(0.0236)	(0.0202)	(0.0194)	(0.0168)
Pop×GDP	0.125***	0.0671***	0.0169	-0.219*	0.0254	0.0100	0.142***	0.0889***	0.0690**
	(0.0384)	(0.0217)	(0.0259)	(0.123)	(0.0240)	(0.0319)	(0.0430)	(0.0342)	(0.0267)
Observations	18,369	18,370	18,370	343	6,521	14,831	18,026	11,849	3,539
County/Mun. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Admin Areas	3,446	3,534	2,933	83	1,103	2,243	3,363	2,431	690

Cluster-robust standard errors in parentheses; clustered at county/município level

*** p<0.01, ** p<0.05, * p<0.1

Table 9: Estimates by Tercile of GDP

	Combined			USA			BRA		
Terciles of Pop	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	2.526***	-1.138	-0.294				2.753***	-0.532	0.234
	(0.906)	(1.391)	(1.233)				(0.926)	(1.536)	(1.324)
GDP	0.674***	1.101**	0.130	0.199	0.0691	0.0711	0.443	1.285**	0.661***
	(0.227)	(0.456)	(0.104)	(0.478)	(0.856)	(0.121)	(0.362)	(0.608)	(0.249)
Pop	1.466	0.655	5.604***	0.363	-9.674**	3.801***	1.149	1.321	4.949***
	(0.955)	(3.516)	(0.709)	(2.949)	(4.614)	(0.867)	(1.092)	(4.648)	(1.382)
GDP^2	-0.0207**	-0.0224*	0.00485***	-0.0122	-0.0365**	0.00509*	-0.0431***	-0.0243	-0.0520***
	(0.00922)	(0.0134)	(0.00188)	(0.0144)	(0.0169)	(0.00275)	(0.0147)	(0.0225)	(0.0121)
Pop^2	-0.00327	0.0200	-0.261***	0.0200	0.382*	-0.261***	-0.0442	-0.0144	-0.290***
	(0.0429)	(0.190)	(0.0386)	(0.162)	(0.224)	(0.0417)	(0.0473)	(0.263)	(0.0722)
$Area^2$	-0.112*	0.116	-0.0565				-0.100	0.118	-0.0616
	(0.0629)	(0.0803)	(0.0811)				(0.0620)	(0.0805)	(0.0823)
Area×Pop	-0.128**	-0.0362	0.104	0.0404	0.380*	0.341***	-0.158***	-0.0705	0.0817
	(0.0563)	(0.120)	(0.0695)	(0.187)	(0.198)	(0.101)	(0.0558)	(0.147)	(0.0978)
Area×GDP	-0.00786	0.0173	0.0115	0.0426	0.144***	0.0176	-0.0121	-0.00798	-0.00539
	(0.0163)	(0.0217)	(0.0109)	(0.0592)	(0.0529)	(0.0181)	(0.0195)	(0.0257)	(0.0149)
Pop×GDP	-0.00900	-0.0577	-0.0280*	-0.0291	-0.0124	-0.0281	0.0852*	-0.0557	0.0852**
	(0.0209)	(0.0647)	(0.0165)	(0.0329)	(0.113)	(0.0241)	(0.0453)	(0.0891)	(0.0368)
Observations	18,382	18,380	18,380	4,366	6,643	10,719	14,016	11,737	7,661
County/Mun. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Admin areas	3,043	3,047	2,861	644	987	1,547	2,399	2,060	1,314

Cluster-robust standard errors in parentheses; clustered at county/município level

*** p<0.01, ** p<0.05, * p<0.1

Table 10: Estimates by Tercile of Population

4.7 GDP×Area

4.7.1 Areal Tercile 1; <467km^2

The next set of tables, tables 11-13, breaks down the areal terciles again, this time with each areal tercile is further broken into terciles of GDP. The following tables will follow a similar format as before: the left three columns are the combined estimates, the center three columns are the estimates from the USA sample alone, and the last three columns are the estimates utilizing only the Brazilian sample. The first table, table 11, is the smallest tercile, with the terciles of area labeled at the bottom of the table, the second is the middle tercile which is administrative areas larger than 400km2 and smaller than 1600km2, and the third tercile is those larger than 1600km2. Even in the smallest category, which does not have statistically significant effects in the combined regressions from earlier, now has a statistically significant, though not at traditional levels, and positive combined effect for the wealthiest tercile of administrative areas. Among the smallest counties, the GDP^2 effect appears to vary between income groups, the effect being negative and large for the wealthier counties and large and negative also for the middle income tercile in the USA sample. Last, with respect to the population*GDP variable, which was negative except for the Brazilian sample in the previous tables, we see the negative relationship is driven by the smallest, poorest tercile of counties in the USA sample. The middle and wealthier terciles are estimated to have a positive relationship meaning that as population increases, the strength of the GDP lights relationship grows stronger even holding GDP constant, and vice versa where as GDP increases, the strength of the population-lights relationship also increases.

4.7.2 Areal Tercile 2; $467km^2 < c/m < 1495km^2$

Looking at the second-largest tercile of administrative areas, table 12, those larger than 400 but less than 1600 sq km, we see that where there was no statistically significant estimate for the GDP variable in the prior combined table, now we can see that for the wealthiest tercile of administrative areas there is still a strong and statistically significant positive relationship between GDP and nighttime light. In the Brazilian sample the coefficient estimates are smaller and not significant at the standard levels. Turning to the estimate for the effect of population on nighttime light, at least for the middle areal tercile we see that the effects are predominantly estimated to be negative across the columns. This could represent the true relationship or be the result of some kind of endogeneity where light precedes an increase in population. The second order terms, GDP appears to be small and negative as in the previous estimates, this parameter appears to be well estimated across many specifications. For the population-squared estimates the effect is positive and significant in some cases, though for the poorest tercile it appears the effect might be negative, though not significant at standard levels.

Looking at the areal interaction terms we see that for both, the USA sample has positive and significant effects meaning that larger counties within this tercile the strength of the GDP nighttime lights relationship is increasing in the area of the county, and the same is true of the population-lights relationship. The last interaction is the population*GDP interaction term.

In previous tables the estimates put this as having strong positive effects for this areal tercile, (column 2, table 5) the poorest tercile from the USA sample has a negative relationship, with the magnitude estimated to be quite strong. The bottom tercile of GDP showing a negative effect is consistent with previous tables.

4.7.3 Areal Tercile 3; $1495km^2 < c/m$

Turning to the final table of the areal tables, table 13, this table comprises the largest counties/municípios. Looking at the GDP row, for the largest tercile of counties the estimates of the effect of GDP on nighttime light are positive. The combined samples estimate large effects for the middle and wealthiest terciles while in the USA sample the wealthiest tercile has an effect magnitude of .782, though only significant at the 10% significance level, and in Brazil the middle GDP tercile has a statistically significant effect on light, with the magnitude of the effect being much larger than the USA effect at 1.623. With respect to the population variable we see that overwhelmingly the effect is positive and statistically significant for the largest tercile. Turning to the second-order terms the GDP^2 term is small and negative, and also statistically significant consistent with the previous estimates. Whereas in the previous tables we saw some heterogenous second-order effects for the population variable, in this table of the largest areal tercile we see that it is negative across the board, with the effect size being consistently larger than the second-order term for GDP. Turning to the areal interaction terms, the GDP*area term is small in magnitude, .0586-.289 for the statistically significant estimates. This implies that for larger counties and municípios *within* the largest tercile which is counties and municípios larger than 1600km.

Terciles of Area	1								
	Combined			USA			BRA		
Terciles of GDP	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	1.382	-3.546	11.47***				1.377	-3.386	11.64***
	(1.707)	(6.979)	(4.427)				(1.707)	(6.995)	(4.471)
GDP	-0.623	-1.110	3.625	-9.104	1.268	3.423	-0.624	-1.076	3.701
	(0.463)	(0.952)	(2.207)		(1.871)	(3.962)	(0.463)	(1.002)	(2.461)
Pop	2.862	-12.72**	-7.859**	-1,763	37.07**	-12.80	2.870	-13.31**	-5.948
	(2.392)	(5.401)	(3.960)		(14.68)	(9.599)	(2.391)	(5.552)	(9.252)
GDP^2	0.0132	0.00816	-0.332**	0.996	0.0205	-0.0315	0.0132	0.00576	-0.339**
	(0.0275)	(0.0537)	(0.133)		(0.0789)	(0.144)	(0.0275)	(0.0540)	(0.138)
Pop^2	-0.146	0.376	0.217	46.00	-0.523	0.960***	-0.147	0.405	0.112
	(0.150)	(0.296)	(0.163)		(0.416)	(0.274)	(0.150)	(0.306)	(0.454)
Area^2	-0.0641	-0.0809	-0.878***				-0.0642	-0.0791	-0.886***
	(0.0643)	(0.424)	(0.287)				(0.0643)	(0.423)	(0.284)
Area×Pop	-0.0744	0.573	0.150	175.8	-3.964***	0.513	-0.0734	0.565	0.141
	(0.206)	(0.587)	(0.253)		(1.318)	(1.308)	(0.206)	(0.591)	(0.265)
Area×GDP	0.0213	-0.0461	0.161	-0.360	-0.0794	-0.0279	0.0211	-0.0558	0.166
	(0.0370)	(0.0978)	(0.128)		(0.102)	(0.176)	(0.0371)	(0.0996)	(0.130)
Pop×GDP	0.0395	0.164*	0.327***	-1.248	-0.159	-0.399**	0.0398	0.174*	0.333**
	(0.0499)	(0.0893)	(0.122)		(0.125)	(0.168)	(0.0499)	(0.0928)	(0.161)
Observations	11,450	4,723	2,196	21	90	232	11,429	4,633	1,964
County/Mun. FE	Yes	Yes	Yes	Yes			Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes			Yes	Yes	Yes
Number of Admin areas	2,095	915	461	3	23	57	2,092	892	404

Cluster-robust standard errors in parentheses; clustered at county/município level
*** p<0.01, ** p<0.05, * p<0.1

Table 11: First Areal Tercile, Counties and Municípios < 467 km^2; By terciles of GDP

Terciles of GDP	2								
	Combined			USA			BRA		
Terciles of Pop	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	-19.60*	3.127	-2.089				-13.63	3.335*	-2.089
	(10.23)	(1.940)	(1.787)				(12.47)	(2.016)	(1.782)
GDP	1.681***	-0.367	0.323	1.096**	-0.566	105.7*	0.407	-0.330	0.334
	(0.534)	(0.823)	(1.095)	(0.478)	(0.719)	(50.27)	(1.628)	(1.019)	(1.092)
Pop	-1.987	-4.301	5.079	-6.236**	-15.25*	21.18	6.589	-2.632	5.002
	(3.157)	(5.099)	(7.512)	(2.805)	(8.502)	(364.9)	(12.66)	(6.245)	(7.501)
GDP^2	-0.0675***	-0.00328	-0.0799**	-0.0689***	-0.0315	-1.992	-0.0487	-0.00895	-0.0798**
	(0.0206)	(0.0270)	(0.0345)	(0.0223)	(0.0277)	(1.591)	(0.0447)	(0.0310)	(0.0344)
Pop^2	0.174	0.256	-0.422	0.290**	0.711**	3.418	-0.342	0.167	-0.418
	(0.157)	(0.274)	(0.394)	(0.137)	(0.336)	(22.59)	(0.782)	(0.354)	(0.394)
Area^2	1.696**	-0.0996	0.0845				1.469**	-0.100	0.0845
	(0.663)	(0.0892)	(0.102)				(0.747)	(0.0884)	(0.102)
Area×Pop	-0.120	-0.180	0.185	0.153	0.283	-1.420	-0.385	-0.197	0.185
	(0.158)	(0.146)	(0.178)	(0.133)	(0.419)	(14.92)	(0.408)	(0.157)	(0.178)
Area×GDP	-0.0304	0.0391	-0.0491*	0.0330	0.246***	0.839	-0.0901	0.0361	-0.0491*
	(0.0194)	(0.0276)	(0.0279)	(0.0207)	(0.0405)	(1.227)	(0.0583)	(0.0306)	(0.0278)
Pop×GDP	0.0211	0.0324	0.215*	0.0378	-0.0405	-5.892	0.153	0.0475	0.214*
	(0.0327)	(0.101)	(0.122)	(0.0343)	(0.0788)	(4.753)	(0.128)	(0.122)	(0.122)
Observations	5,269	8,922	4,179	3,596	2,874	51	1,673	6,048	4,128
County/Mun. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Admin area	1,029	1,817	840	583	538	12	446	1,279	828

Cluster-robust standard errors in parentheses; clustered at county/município level

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 12: Second Areal Tercile, 467 $\text{km}^2 \leq$ Counties and Municípios $\leq$1400 km^2; By terciles of GDP

Terciles of Area	3								
	Combined			USA			BRA		
Terciles of GDP	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	-4.092*	12.67	5.550**				-4.367*	4.378	8.594***
	(2.370)	(8.847)	(2.296)				(2.409)	(9.615)	(3.229)
GDP	0.799*	0.787**	0.952***	1.795**	0.0332	0.919**	0.660	1.903***	0.822
	(0.432)	(0.376)	(0.363)	(0.699)	(0.410)	(0.414)	(0.468)	(0.730)	(1.479)
Pop	4.169**	3.484*	4.241***	-0.276	7.295***	2.953**	4.782**	-5.193	7.464
	(1.888)	(2.050)	(1.303)	(2.373)	(2.089)	(1.386)	(2.096)	(4.823)	(7.043)
GDP^2	-0.0489***	-0.0717***	-0.0447***	-0.0512	-0.0702***	-0.0463***	-0.0499***	-0.0991***	-0.0545
	(0.0157)	(0.0118)	(0.0157)	(0.0407)	(0.0130)	(0.0163)	(0.0156)	(0.0256)	(0.0647)
Pop^2	-0.267***	-0.129***	-0.178***	-0.127	-0.165***	-0.177***	-0.294***	-0.0207	-0.207
	(0.0728)	(0.0446)	(0.0513)	(0.0855)	(0.0443)	(0.0526)	(0.0829)	(0.138)	(0.309)
$Area^2$	0.0248	-0.785	-0.323**				0.0347	-0.758	-0.176
	(0.150)	(0.617)	(0.130)				(0.138)	(0.629)	(0.198)
Area×Pop	0.313*	-0.212	-0.0605	0.476*	-0.634**	0.103	0.319	0.613	-0.472
	(0.186)	(0.247)	(0.114)	(0.263)	(0.293)	(0.125)	(0.202)	(0.431)	(0.300)
Area×GDP	0.0435	0.109***	0.0469	-0.0667	0.218***	0.0617	0.0493	-0.00865	-0.0180
	(0.0419)	(0.0400)	(0.0329)	(0.0504)	(0.0542)	(0.0380)	(0.0444)	(0.0691)	(0.0610)
Pop×GDP	0.0441	0.0539**	-0.00175	0.0189	0.0489	-0.00597	0.0561	0.0996**	0.0888
	(0.0334)	(0.0260)	(0.0357)	(0.0658)	(0.0325)	(0.0376)	(0.0355)	(0.0435)	(0.146)
Observations	1,925	6,976	9,469	367	5,806	8,658	1,558	1,170	811
County/Mun. FE	Yes	Yes	Yes	Yes			Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes			Yes	Yes	Yes
Number of Admin areas	357	1,111	1,467	56	882	1,305	301	229	162

Cluster-robust standard errors in parentheses; clustered at county/município level
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Third Areal Tercile, 1400 km^2 < Counties and Municípios ; By terciles of GDP

4.8 Pop×GDP Terciles

4.8.1 GDP Tercile 1 ; T <$61,583

The next tables, 14-16, following a similar format, decompose the effects of population into different terciles, by tercile of GDP. The first table, table 14, is the lowest tercile of GDP and represents the poorest counties. As we can see there are far more poor municipalities/counties in Brazil than there are in the USA. The poorest tercile yields some interesting results, some of which are not subtle. Looking across the GDP row we see some wildly different estimates for the effect of GDP on nighttime light. At least in the smallest population tercile (administrative areas $< 10,000$ persons) we see that the GDP^2 term is estimated to be fairly consistent with all the past estimates driving home the point about endogeneity. Also of note, the population×GDP effect is positive and statistically significant in the lowest tercile, whereas, at least in the combined estimates, the estimate for the most populated tercile is strong and negative meaning that, in the most populous counties and municípios, the effect of higher income leads to a sharply weaker relationship between GDP and light, or a higher population leads to a sharply weaker relationship between Population and lights.

4.8.2 GDP Tercile 2 ; $61,583 < T < $438,452

The second table in this section, table 15, corresponds to the second tercile of GDP, which is counties with GDP greater than $61,583 and less than $438,452. checking the first row of interest, the GDP effect, we see fairly stable estimates, thought only the middle population tercile for the USA sample has a significant effect. Interestingly for the middle-GDP tercile and the USA sample the effect of increases in population appear to decrease light, which holds for the two lowest terciles of population. The second order terms, GDP^2 appears to be well-estimated again similar to all of the last tables, with the estimates statistically significant in the lowest population tercile and the largest, at the 10% and 5% level respectively. In the USA sample the middle tercile of population has a statistically significant and negative second-order effect of .101. For the population second-order terms, they are positive and, in the bottom two terciles of the USA sample, statistically significant at standard levels. For the rest of this table there is not much with respect to the areal interaction terms to mention except in column two of the USA sample, the area*GDP effect is statistically significant and positive meaning larger counties and municípios within this category experience a stronger GDP-lights relationship.

4.8.3 GDP Tercile 3 ; $438,452 < T

The final table in the sub-sample analysis, table 16, is the top tercile of GDP which is the most productive counties and municípios, broken down by terciles of population. A large sample is not available for the lowest population tercile, counties and municípios smaller than 10,000, only 90 counties or municṕios fall into this category. The effect of GDP on light is estimated to be negative in this category, though positive in the most populated tercile with respect to the combined estimates. This pattern holds across the USA and Brazilian samples as well. For the population variable in the next row after GDP there is actually a similar pattern, with the effects for the least populated counties and municṕios being negative while for the more populated (over 25,000 persons) areas light is strongly increasing in population. Turning to the squared terms, for the GDP^2 term, the parameter estimate is positive and fairly large for the least populated areas, and negative and modest in magnitude for the most populated areas. Looking at the $population^2$ term the effect is estimated to be negative and fairly large for the top tercile of populated counties and municípios while for the rest of the columns none of the estimates are statistically significant at standard levels. Looking at the areal interaction terms, for both population*area and GDP*area the effect is estimated to be large and statistically significant in the wealthiest counties in the USA. In the last row, the population*GDP interaction effect the estimates vary drastically and the only column which has a statistically significant effect is the top tercile of the Brazilian sample.

Terciles of GDP	1								
	Combined			USA			BRA		
Terciles of Pop	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	2.539*	-2.593	-23.68				2.428*	-2.593	-23.68
	(1.327)	(2.070)	(27.35)				(1.439)	(2.070)	(27.21)
GDP	0.551	2.186*	-57.84*	2.624	-	-	0.503	2.186*	-57.84*
	(0.540)	(1.253)	(32.18)	(2.476)	-	-	(0.598)	(1.253)	(32.03)
Pop	-2.670	-3.596	1,129***	-12.22**	-	-	-1.424	-3.596	1,129***
	(2.452)	(8.291)	(273.2)	(5.739)	-	-	(4.179)	(8.291)	(272.3)
GDP^2	-0.0514*	-0.110	0.307	-0.0626	-	-	-0.0522*	-0.110	0.307
	(0.0298)	(0.0707)	(1.039)	(0.0908)	-	-	(0.0300)	(0.0707)	(1.034)
Pop^2	0.150	-0.000912	-59.46***	0.795***	-	-	0.0655	-0.000912	-59.46***
	(0.151)	(0.479)	(13.03)	(0.244)	-	-	(0.284)	(0.479)	(12.99)
$Area^2$	-0.133**	0.0427	-1.121				-0.131**	0.0427	-1.121
	(0.0661)	(0.0928)	(0.987)				(0.0654)	(0.0928)	(0.982)
Area×Pop	-0.0963	0.355	4.417*	0.479			-0.0831	0.355	4.417*
	(0.135)	(0.231)	(2.421)	(0.599)			(0.156)	(0.231)	(2.409)
Area×GDP	-0.0126	-0.0942**	-0.508	0.0135			-0.0132	-0.0942**	-0.508
	(0.0225)	(0.0374)	(0.639)	(0.0564)			(0.0248)	(0.0374)	(0.636)
Pop×GDP	0.0970**	0.121	5.283**	-0.219*			0.106*	0.121	5.283**
	(0.0483)	(0.182)	(2.612)	(0.122)			(0.0597)	(0.182)	(2.601)
Observations	12,673	5,531	165	340			12,333	5,531	162
County/Mun. FE	Yes	Yes	Yes	Yes			Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes			Yes	Yes	Yes
Number of Admin areas	2,304	1,148	68	82			2,222	1,148	67

Cluster-Robust standard errors in parentheses; county/municṕio level
*** p<0.01, ** p<0.05, * p<0.1

Table 14: GDP Tercile 1, Counties/Municípios <143,217$; By tercile of population

Terciles of GDP	2								
	Combined			USA			BRA		
Terciles of Pop	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	-19.60*	3.127	-2.089				-13.63	3.335*	-2.089
	(10.23)	(1.940)	(1.787)				(12.47)	(2.016)	(1.782)
GDP	1.681***	-0.367	0.323	1.096**	-0.566	105.7*	0.407	-0.330	0.334
	(0.534)	(0.823)	(1.095)	(0.478)	(0.719)	(50.27)	(1.628)	(1.019)	(1.092)
Pop	-1.987	-4.301	5.079	-6.236**	-15.25*	21.18	6.589	-2.632	5.002
	(3.157)	(5.099)	(7.512)	(2.805)	(8.502)	(364.9)	(12.66)	(6.245)	(7.501)
GDP^2	-0.0675***	-0.00328	-0.0799**	-0.0689***	-0.0315	-1.992	-0.0487	-0.00895	-0.0798**
	(0.0206)	(0.0270)	(0.0345)	(0.0223)	(0.0277)	(1.591)	(0.0447)	(0.0310)	(0.0344)
Pop^2	0.174	0.256	-0.422	0.290**	0.711**	3.418	-0.342	0.167	-0.418
	(0.157)	(0.274)	(0.394)	(0.137)	(0.336)	(22.59)	(0.782)	(0.354)	(0.394)
$Area^2$	1.696**	-0.0996	0.0845				1.469**	-0.100	0.0845
	(0.663)	(0.0892)	(0.102)				(0.747)	(0.0884)	(0.102)
Area×Pop	-0.120	-0.180	0.185	0.153	0.283	-1.420	-0.385	-0.197	0.185
	(0.158)	(0.146)	(0.178)	(0.133)	(0.419)	(14.92)	(0.408)	(0.157)	(0.178)
Area×GDP	-0.0304	0.0391	-0.0491*	0.0330	0.246***	0.839	-0.0901	0.0361	-0.0491*
	(0.0194)	(0.0276)	(0.0279)	(0.0207)	(0.0405)	(1.227)	(0.0583)	(0.0306)	(0.0278)
Pop×GDP	0.0211	0.0324	0.215*	0.0378	-0.0405	-5.892	0.153	0.0475	0.214*
	(0.0327)	(0.101)	(0.122)	(0.0343)	(0.0788)	(4.753)	(0.128)	(0.122)	(0.122)
Observations	5,269	8,922	4,179	3,596	2,874	51	1,673	6,048	4,128
County/Mun. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Admin area	1,029	1,817	840	583	538	12	446	1,279	828

Cluster-Robust standard errors in parentheses; county/municṕio level

*** p<0.01, ** p<0.05, * p<0.1

Table 15: GDP Tercile 2 , \$143,217 < counties/municṕios < \$639,889; By tercile of population

Terciles of GDP	3								
	Combined			USA			BRA		
Terciles of Pop	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area		50.44	0.401					51.67	0.773
		(39.34)	(1.123)					(47.24)	(1.281)
GDP	-2.068	0.551	0.969***	-2.068	0.534	0.692**	8.273	0.0489	0.484
	(3.228)	(1.274)	(0.204)	(3.212)	(1.340)	(0.326)	(801,564)	(9.770)	(0.297)
Pop	-13.75	-2.626	4.197***	-13.75	-5.263	2.550**	-98,423	-30.06	6.898***
	(10.63)	(6.747)	(0.761)	(10.58)	(7.570)	(0.991)	(4.201e+08)	(34.24)	(1.636)
GDP^2	0.159	-0.0358*	-0.0317*	0.159	-0.0381*	-0.0184	-0.287	0.0260	-0.0509***
	(0.101)	(0.0200)	(0.0179)	(0.100)	(0.0212)	(0.0380)	(27,865)	(0.271)	(0.0146)
Pop^2	0.241	0.142	-0.211***	0.241	0.211	-0.168***	6,005	1.850	-0.380***
	(0.464)	(0.309)	(0.0465)	(0.462)	(0.326)	(0.0648)	(2.562e+07)	(1.912)	(0.0845)
$Area^2$		-2.923	-0.165***					-2.404	-0.167***
		(2.322)	(0.0619)					(2.737)	(0.0641)
Area×Pop	1.298	-0.0399	0.132**	1.298	0.139	0.330***	-860.0	-1.003**	0.129
	(1.230)	(0.197)	(0.0635)	(1.224)	(0.229)	(0.106)	(3.639e+06)	(0.394)	(0.0891)
Area×GDP	-0.371	0.0231	0.0290**	-0.371	0.0356	0.0955***		-0.0327	0.00629
	(0.227)	(0.0515)	(0.0142)	(0.226)	(0.0581)	(0.0275)		(0.161)	(0.0158)
Pop×GDP	0.0455	0.0316	-0.0173	0.0455	0.0296	-0.0735		-0.0558	0.0916***
	(0.159)	(0.136)	(0.0438)	(0.158)	(0.141)	(0.0916)		(0.379)	(0.0354)
Observations	440	3,925	14,005	430	3,767	10,634	10	158	3,371
County/Mun. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of new_id	108	696	2,179	104	644	1,538	4	52	641

Cluster-Robust standard errors in parentheses; county/municṕio level
*** p<0.01, ** p<0.05, * p<0.1

Table 16: GDP Tercile 3 ,$639,889 < counties/municṕios; By tercile of population

4.9 Regression Using Geographic Characteristics

4.9.1 Between-county Estimator

In table 17, the first set of geographic regressions follows the same translog specification as before. The difference is that now the counties and municípios have been collapsed to their mean values, and thus a normal OLS regression corresponds to the between-county or between-município estimates. In this context, we are able to estimate the effects of time-invariant features of certain counties on overall light. The implied comparison here is other counties which do not have airports, primary roads, railways, ports, or border crossings. Although this procedure does benefit in that it allows us to estimate the effects, it does not rule out the possibility that some other omitted variable may be simultaneously determining both infrastructure and light (such as the presence of a mountain), and infrastructure elements, especially ports and roads, are rarely placed randomly. Looking at the GDP row, the effects of GDP for the between estimates are estimated to be quite large and around 1.1-1.2 in magnitude. Strangely, restricted to the USA sample the effects of GDP on light are then negative, while in the Brazilian sample they remain positive. The changing of the sign on the USA sample might be related to the endogeneity of population to GDP and light. Light is also estimated to be increasing in population, though less strongly than it is with GDP, with the estimates of the relationship much larger for the USA sample than for the Brazilian sample. The GDP^2 term is negative, fairly small in magnitude and statistically significant consistent with many of the previous estimates. In the case of the USA sample the sign changes to positive, indicating that, at least in some counties, there are increasing returns to GDP with respect to light. For $population^2$ almost the inverse is true, the combined effect is estimated to be positive, the USA-restricted sample is also positive, while in the Brazilian sample there are negative effects meaning diminishing returns to population's effect on nighttime light. The areal interaction terms, Population*area and GDP*area, in the case of area*pop the relationship is negative and sizeable, though the effect is estimated to be much smaller for Brazil. The area*GDP effect is estimated to be positive, and is significant across all columns; the combined estimates are larger than the restricted estimates, where the restricted estimates put the magnitude of this relationship between .052 and .085. Put differently, larger counties and municípios have a stronger relationship between GDP and light, which is consistent with some of the previous tables. Next looking at the population*GDP interaction term we see that the effect is not statistically significant for the combined estimates, though for the USA-restricted sample the estimated effect is negative, while for the Brazilian sample the effect is small and positive.

Finally reviewing the indicator variables for geographic features, the first variable is the port variable which takes 1 if a county or município has a port and 0 otherwise and is thus the marginal effect of having a port on nighttime light. This effect is estimated to be 0.181 meaning having a port increases light in the period 2012-2017 by .18 percent. The effect is statistically significant in the combined sample and the USA sample, though in the Brazilian sample the effect size appears to be much smaller and is no longer statistically significant. Strangely for the next row, the indicator for the presence of a primary road, the overall effect is estimated to

VARIABLES	(1) NTL	(2) NTL	(3) NTL	(4) NTL	(5) NTL	(6) NTL
Area	0.306***	0.289***	0.495***	0.489***	-0.142**	-0.142**
	(0.0804)	(0.0776)	(0.173)	(0.179)	(0.0688)	(0.0692)
GDP	1.182***	1.203***	-0.808***	-0.801***	0.812***	0.812***
	(0.126)	(0.123)	(0.252)	(0.269)	(0.109)	(0.107)
Pop	0.392**	0.401***	1.376***	1.377***	0.308**	0.307**
	(0.158)	(0.153)	(0.246)	(0.267)	(0.129)	(0.127)
GDP^2	-0.0621***	-0.0630***	0.0714***	0.0711***	-0.103***	-0.103***
	(0.0145)	(0.0147)	(0.0182)	(0.0193)	(0.0118)	(0.0116)
Pop^2	0.0831***	0.0834***	0.102***	0.102***	-0.118***	-0.118***
	(0.0198)	(0.0198)	(0.0168)	(0.0176)	(0.0184)	(0.0186)
$Area^2$	-0.000670	-0.000271	0.0535***	0.0538***	0.00603	0.00607
	(0.00566)	(0.00583)	(0.0122)	(0.0127)	(0.00394)	(0.00390)
Area×Pop	-0.282***	-0.282***	-0.189***	-0.189***	-0.0639***	-0.0640***
	(0.0217)	(0.0217)	(0.0345)	(0.0341)	(0.0124)	(0.0128)
Area×GDP	0.201***	0.202***	0.0844***	0.0845***	0.0525***	0.0526***
	(0.0165)	(0.0162)	(0.0324)	(0.0322)	(0.00980)	(0.00991)
Pop×GDP	-0.0113	-0.0121	-0.141***	-0.141***	0.215***	0.215***
	(0.0324)	(0.0331)	(0.0295)	(0.0319)	(0.0276)	(0.0275)
Has Port	0.181***	0.144**	0.230***	0.223***	0.0599	0.0705
	(0.0576)	(0.0602)	(0.0516)	(0.0516)	(0.111)	(0.127)
Has Road	-0.682***	-0.689***	0.103***	0.101***	0.354***	0.354***
	(0.0250)	(0.0244)	(0.0139)	(0.0145)	(0.0471)	(0.0484)
Has Rail	0.539***	0.538***	0.0854***	0.0854***	0.0861***	0.0863***
	(0.0229)	(0.0220)	(0.0240)	(0.0234)	(0.0170)	(0.0169)
Has Border Crossing	0.392***	0.388***	0.0355	0.0345	0.151	0.150
	(0.104)	(0.106)	(0.0694)	(0.0678)	(0.107)	(0.110)
Has Airport	0.291***	0.268***	0.0185	0.0142	-0.0938*	-0.0925*
	(0.0304)	(0.0325)	(0.0188)	(0.0201)	(0.0497)	(0.0516)
Has All Four		0.147***		0.0262		-0.0545
		(0.0377)		(0.0343)		(0.234)
Observations	8,671	8,671	3,101	3,101	5,570	5,570
R-squared	0.863	0.863	0.894	0.894	0.832	0.832
Number of new_id	8,671	8,671	3,101	3,101	5,570	5,570

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 17: Between-county Estimates

be negative while the effect in the two sub-samples is estimated to be positive and statistically significant. This could owe to the fact that the placement of roads is highly endogenous and thus we have unreliable point-estimates, though it seems highly unlikely that having a primary road would lead to having less light. One of the most interesting geographic features driving light is the presence of railway infrastructure which increases light substantially in the combined estimates, with a smaller effect of rail infrastructure on light estimated for the sub-samples of the USA and Brazil. The effect of a border crossing point is next, and these are all the land crossing points captured in the dataset. The combined estimates indicate that these contribute significantly to light, and although the effect is positive for all columns, it is only statistically significant for the combined estimates. A similar story is true for the airport variable, which has positive and statistically significant effects in the combined estimates though not for the individual country estimates, which are actually negative for the Brazilian sample. This may owe to the fact that there are more rural, less active airports in Brazil. Last, an interaction term is included for counties and municípios which contain an airport, a major road, railway, and a port and the marginal effect of having these combined is shown to be statistically significant, and large in magnitude, though only significant for the combined estimates and not for the disaggregated samples.

4.10 Model with Long Differences, 2017-2012

To compare to the previous work on nighttime lights from Henderson et al. (2012) I also estimate a model using the same technique of taking long-differences of each variable. These estimates can be found in the next table of geographic regression variables, table 18. Dummy variables are also included for time-invariant infrastructure characteristics, and I am therefore empirically testing for the significance of different intercepts for each of those categories: roads, rail, ports, crossings, and airports. The first row showing the effect of a change in GDP on the change in nighttime light we see the effect for the combined sample is estimated to be medium-sized in magnitude and negative as well as statistically significant. Much of the effect in the combined estimate appears to be driven by the Brazilian sample which finds similar-sized negative effects of changes in GDP on changes in nighttime light. The effect of population, in the long run, is overwhelmingly positive and significant at the highest levels and across all columns. None of the second-order terms is statistically significant. Moving to the geographic features, starting with ports and including all of the indicators such as road, rail, border crossings and airports, as well as the interaction terms that are equal to one if the county or município contains roads, rails, airports, and a port together. Ports, railways, and airports appear to be the largest contributors to lights according to the combined estimates. The effect of roads is interestingly estimated to be negative using the combined sample, but positive and statistically significant in both of the divided samples. The marginal effect of the interaction term (which takes 1 if a county has 4 of the infrastructure components together) is also positive and statistically significant, large for the combined sample and slightly smaller though still significant for the USA sample. The standard errors in the Brazilian sample for this variable are large and the effect size small, the

VARIABLES	(1) NTL	(2) NTL	(3) NTL	(4) NTL	(5) NTL	(6) NTL
GDP	-0.285**	-0.266**	0.0376	0.0380	-0.241**	-0.240***
	(0.114)	(0.117)	(0.100)	(0.0973)	(0.0936)	(0.0929)
Pop	3.065***	3.248***	4.993***	4.950***	7.375***	7.365***
	(0.324)	(0.313)	(0.372)	(0.363)	(0.550)	(0.567)
Area	-0.930	-3.302			-9.973*	-9.998*
	(0.640)	(4.282)			(5.785)	(5.774)
GDP^2	-0.00760	-0.00529	0.0104	0.0107	0.0159	0.0150
	(0.0572)	(0.0593)	(0.0474)	(0.0468)	(0.0582)	(0.0566)
Pop^2	-0.509	-0.515	1.454	1.670	-1.152	-1.151
	(0.757)	(0.559)	(4.118)	(4.113)	(1.127)	(1.037)
$Area^2$	1.125	0.486			0.686	0.679
	(1.407)	(1.161)			(1.066)	(1.096)
Area×Pop		0.233			0.888	0.891
		(0.438)			(0.605)	(0.603)
Area×GDP	-0.810	-0.409			-0.581	-0.574
	(1.713)	(0.968)			(1.155)	(1.148)
Pop×GDP	0.545	0.321	-0.0132	-0.0239	0.514	0.513
	(0.661)	(0.519)	(0.757)	(0.754)	(0.970)	(0.872)
Has Port	1.124***	0.904***	0.858***	0.743***	0.979***	1.084***
	(0.117)	(0.121)	(0.103)	(0.106)	(0.268)	(0.272)
Has Road	-0.666***	-0.717***	0.481***	0.442***	1.156***	1.155***
	(0.0410)	(0.0406)	(0.0271)	(0.0272)	(0.0680)	(0.0667)
Has Rail	1.762***	1.744***	0.334***	0.334***	0.912***	0.913***
	(0.0359)	(0.0354)	(0.0517)	(0.0520)	(0.0392)	(0.0408)
Has Border Crossing	0.512***	0.508***	0.306***	0.304***	0.803***	0.797***
	(0.118)	(0.117)	(0.106)	(0.106)	(0.296)	(0.292)
Has Airport	1.824***	1.676***	0.911***	0.842***	2.035***	2.053***
	(0.0436)	(0.0461)	(0.0349)	(0.0376)	(0.121)	(0.128)
Has All Four		0.740***		0.344***		-0.550
		(0.0742)		(0.0651)		(0.909)
Observations	8,675	8,675	3,104	3,104	5,571	5,571
R-squared	0.545	0.548	0.495	0.501	0.303	0.303

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18: Long-Differenced Model

estimates are therefore estimated to be negative but they are not significant.

4.11 Placebo Test

As a test for parameter stability, although as we have seen there are some inconsistent results for different parts of the distribution, I drop sequentially one year's worth of data from the sample, and repeat the same regressions. This is akin to a jackknife procedure, and in this case I am using it to confirm the global estimates. The results for these tests are shown in the appendix table 24. All parameter estimates appear to be stable despite the dropping of a year's worth of data. If the effect of GDP on nighttime light were poorly estimated we would see a large variance or potentially changing of the sign on the estimates for the direct effect of GDP on nighttime light.

5 Conclusion

Using quality nationwide panel data from the USA and Brazil, pairing these data with the newest VIIRS night-time satellite imagery, I analyzed the relationship between population, income, geographic variables, and human-generated night-time light measured at the county level. I find that the relationship between nighttime lights, GDP and population changes is strong. These results hold even after incorporating higher-order terms and interaction terms to account for the potential for nonlinearities in the lights-income-population nexus. Decreasing returns to GDP and Population in nighttime light were estimated and confirmed to be present. I also discuss the value-added of nighttime lights over electrical consumption data, and find that electrical consumption is more sensitive to changes in population growth than changes in income. Nighttime light data is available at a monthly frequency and therefore nighhttime lights may be at least as good in place of other data.

I also utilize a between-county estimator to measure the effects of important infrastructure elements on light; infrastructure elements which drive commerce such as roads, rail, ports, and airports are found to substantially influence light production. These findings could be useful to future researchers looking to use VIIRS imagery for economic analysis, for nowcasting small areal GDP, or for policymakers who may be looking to monitor changes in light on a higher-frequency basis. I argue that based on these results, night-time light is found to be a strong proxy indicator for population changes, and a useful indicator for changes in income, though particular attention should be paid to incorporating nonlinear terms.

References

Aysheshim, K., Hinson, J. R., and Panek, S. D. (2020). A primer on local area gross domestic product methodology. *Survey of Current Business*, 100(3):1–13.

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *The quarterly journal of economics*, 127(2):587–644.

Carlowicz, M. (2012). Out of the blue and into the black: New views of the earth at night. [Online; posted 5-December-2012].

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

Donaldson, D. and Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–98.

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

Frick, S. A., Rodríguez-Pose, A., and Wong, M. D. (2019). Toward economically dynamic special economic zones in emerging countries. *Economic Geography*, 95(1):30–64.

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

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

Hu, Y. and Yao, J. (2019). Illuminating economic growth. Technical report, International Monetary Fund.

Kocornik-Mina, A., McDermott, T. K., Michaels, G., and Rauch, F. (2020). Flooded cities. *American Economic Journal: Applied Economics*, 12(2):35–66.

Levin, N. and Zhang, Q. (2017). A global analysis of factors controlling viirs nighttime light levels from densely populated areas. *Remote sensing of environment*, 190:366–382.

Mellander, C., Lobo, J., Stolarick, K., and Matheson, Z. (2015). Night-time light data: A good proxy measure for economic activity? *PloS one*, 10(10).

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

Table 19: Descriptive Statistics for All Regression Variables

		(1) N	(2) mean	(3) sd	(4) min	(5) max	(13) p10	(14) p25	(15) p50	(16) p75	(17) p90
Combined	Total Nighttime Light (Sum of all px)	55,155	7829	43155	0	2922000	110	272	1388	5653	14668
	BLS/IBGE GDP	55,110	2799000	17400000	-19046	710900000	48013	101039	285054	1040000	3919000
	ORNL LandScan Pop.	55,143	48522	222678	18	10140000	1125	2748	7866	24473	81195
	ACS/IBGE Pop.	55,143	63126	269040	14.34	12110000	3574	6733	15507	37720	110326
	Area (km2)	55,155	2110	7482	3.565	380898	152.1	319	949	1865	3687
	Has Port	55,160	0.0139	0.117	0	1	0	0	0	0	0
	Has Rail	55,160	0.479	0.5	0	1	0	0	0	1	1
	Has Road	55,160	0.763	0.425	0	1	0	1	1	1	1
	Has Airport	55,160	0.139	0.346	0	1	0	0	0	0	1
	Has all four	55,160	0.0314	0.174	0	1	0	0	0	0	0
	Has Border Crossing	55,160	0.00988	0.0989	0	1	0	0	0	0	0
USA	Total Nighttime Light (Sum of all px)	21,728	17485	66982	447	2922000	2292	3590	6476	13506	31997
	BLS/IBGE GDP	21,695	5506000	24250000	2753	710900000	162875	335188	874434	2600000	9119000
	ORNL LandScan Pop.	21,728	103045	333748	81	10140000	4821	10569	24921	67781	205340
	ACS/IBGE Pop.	21,728	104246	332430	86	10120000	5144	11021	26017	68958	208518
	Area (km2)	21,728	3004	9610	40.57	380898	806.9	1149	1648	2461	4880
	Has Port	21,728	0.0271	0.162	0	1	0	0	0	0	0
	Has Rail	21,728	0.881	0.324	0	1	0	1	1	1	1
	Has Road	21,728	0.45	0.498	0	1	0	0	0	1	1
	Has Airport	21,728	0.316	0.465	0	1	0	0	0	1	1
	Has all four	21,728	0.078	0.268	0	1	0	0	0	0	0
	Has Border Crossing	21,728	0.019	0.137	0	1	0	0	0	0	0
Brazil	Total Nighttime Light (Sum of all px)	33,427	1553	7530	0	341499	80	154	364	1001	2857
	BLS/IBGE GDP	33,415	1041000	10480000	-19046	699300000	38403	65778	145453	391660	1270000
	ORNL LandScan Pop.	33,415	13068	78808	18	4925000	788	1737	4074	9140	20674
	ACS/IBGE Pop.	33,415	36387	213958	14.34	12110000	3245	5417	11432	24762	56962
	Area (km2)	33,427	1529	5610	3.565	159533	113.8	204.3	417.8	1028	2747
	Has Port	33,432	0.00538	0.0732	0	1	0	0	0	0	0
	Has Rail	33,432	0.218	0.413	0	1	0	0	0	0	1
	Has Road	33,432	0.966	0.181	0	1	1	1	1	1	1
	Has Airport	33,432	0.0244	0.154	0	1	0	0	0	0	0
	Has all four	33,432	0.00395	0.0627	0	1	0	0	0	0	0
	Has Border Crossing	33,432	0.00108	0.0328	0	1	0	0	0	0	0

State	County	year	Total NTL	BLS GDP	LS Pop	ACS Pop	square_miles	square_km
Alaska	Yukon-Koyukuk	2017	2921585	258303	5366	5396	147066	380898
Alaska	Yukon-Koyukuk	2016	2741543	260813	4795	5423	147066	380898
Alaska	Yukon-Koyukuk	2015	2596611	247510	6657	5466	147066	380898
Alaska	Yukon-Koyukuk	2014	2470665	226243	6693	5464	147066	380898
Alaska	Yukon-Koyukuk	2013	2123825	277385	6840	5564	147066	380898
Alaska	North Slope	2017	1989463	11231169	8976	9831	90793	235153
Alaska	North Slope	2015	1941614	11130682	9379	9795	90793	235153
Alaska	Yukon-Koyukuk	2012	1937930	316396	6834	5624	147066	380898
Alaska	North Slope	2016	1867156	10567213	8218	9718	90793	235153
Alaska	North Slope	2018	1769743	10469543	14320	9872	90793	235153
Alaska	North Slope	2013	1620345	7251453	9388	9786	90793	235153
Alaska	North Slope	2012	1131531	8920976	9343	9692	90793	235153
Alaska	Northwest Arctic	2016	980246	591812	6639	7689	36771	95236
Alaska	Northwest Arctic	2017	925620	680814	7527	7767	36771	95236
Alaska	Northwest Arctic	2013	867246	667707	7685	7725	36771	95236
Texas	Harris	2017	824801	351838304	4844329	4664159	1760	4557
California	Los Angeles	2017	822111	688661568	10132862	10118759	4088	10587
Alaska	Northwest Arctic	2015	811720	577594	7719	7771	36771	95236
Texas	Harris	2013	800395	390463008	4472666	4355158	1760	4557
Texas	Harris	2015	783815	358868384	4676992	4561939	1760	4557
Texas	Harris	2014	779031	392944160	4581052	4458709	1760	4557
California	Los Angeles	2018	757890	710893248	10100543	10105518	4088	10587
California	Los Angeles	2014	747704	630438080	10081448	10048408	4088	10587
Illinois	Cook	2014	743964	350384992	5403468	5257481	962	2492
California	Los Angeles	2015	739414	653885056	10143410	10097037	4088	10587
Alaska	Southeast Fairbanks	2017	735827	640754	6888	6885	26183	67813

Table 20: Top 25 US Counties in Total Light 2012-2018

name_1	name_2	year	Total NTL	BLS GDP	LS Pop	ACS Pop	square_miles	square_km
Kentucky	Robertson	2016	447	26076	1984	2125	101	261
Kentucky	Robertson	2012	459	19574	1867	2216	101	261
Washington	Wahkiakum	2016	515	96746	3414	4167	262	678
Kentucky	Robertson	2013	515	19937	1868	2216	101	261
Kentucky	Robertson	2015	524	24690	1791	2135	101	261
Washington	Wahkiakum	2013	528	64330	3583	4033	262	678
Massachusetts	Nantucket	2016	528	1695910	11101	11124	48	126
Virginia	Highland	2016	533	101481	1918	2209	420	1087
Massachusetts	Nantucket	2013	563	1031003	10910	10567	48	126
Washington	Wahkiakum	2015	564	97635	3586	4027	262	678
Massachusetts	Nantucket	2018	576	1791518	11358	11327	48	126
Massachusetts	Nantucket	2014	594	1116569	11352	10839	48	126
Virginia	Rappahannock	2016	598	267250	6420	7352	265	688
Washington	San Juan	2012	599	492193	14860	15849	181	470
Virginia	Mathews	2016	607	174844	6791	8789	89	231
Georgia	Taliaferro	2016	608	40701	1364	1613	195	506
Washington	San Juan	2015	611	601531	15243	16198	181	470
Massachusetts	Dukes	2016	611	1678037	16831	17316	110	286
Massachusetts	Nantucket	2015	613	1673678	11467	10945	48	126
West Virginia	Wirt	2016	622	58728	5165	5767	232	600
Virginia	Highland	2012	633	46315	1767	2234	420	1087
Massachusetts	Nantucket	2017	633	1722140	11411	11270	48	126
Kentucky	Robertson	2018	638	25531	1804	2135	101	261
Georgia	Glascock	2016	644	45753	2680	2979	144	374
Kentucky	Owsley	2016	648	51987	4396	4473	198	513
Washington	San Juan	2016	658	621278	14145	16304	181	470

Table 21: Bottom 25 US Counties in Total Light 2012-2018

State	Município	year	ntl	gdp	LandScan Pop	IBGE Pop	Area km2
RR	Bonfim	2015	0	224232	2099	11739	8095
RR	Mucajaí	2015	0	248327	8046	16380	12461
RR	Alto Alegre	2015	0	221320	4776	16176	25567
AP	Ferreira Gomes	2015	0	351803	622	6901	4974
AP	Pracuúba	2015	0	56518	314	4531	4948
AP	Calçoene	2015	0	136608	365	10163	14232
RR	Caroebe	2015	0	142421	2232	9165	12066
AP	Amapá	2015	0	131867	3027	8622	9168
RR	Boa Vista	2015	0	7581092	89358	320714	5687
AP	Itaubal	2015	0	57149	2885	4949	1623
AP	Serra do Navio	2015	0	60383	283	4938	7713
AP	Cutias	2015	0	64196	834	5407	2179
RR	Iracema	2015	0	126537	2849	10320	14410
AP	Porto Grande	2015	0	295789	2987	19669	4425
RR	São Luiz	2015	0	100434	1336	7407	1527
RR	Caracaraí	2015	0	307049	4078	20261	47409
RR	São João da Baliza	2015	0	124280	3700	7516	4284
AP	Tartarugalzinho	2015	0	165606	2260	15212	6685
AP	Oiapoque	2015	0	305452	5288	24263	22625
RR	Amajari	2015	0	123154	3598	11006	28472
RR	Normandia	2015	0	123235	4117	10148	6967
RR	Cantá	2015	0	209781	3516	16149	7665
RR	Uiramutã	2015	0	97451	2264	9488	8066
AP	Pedra Branca do Amapari	2015	0	288571	2537	13988	9625
RR	Pacaraima	2015	0	145930	2772	11908	8028

Table 22: Top 25 Darkest Counties, Brazil 2012-2017

State	Município	year	ntl	gdp	ls_pop	pop	area
SP	São Paulo	2014	341499	621900000	4248387	11895893	1521
SP	São Paulo	2016	325241	683100000	4312434	12038175	1521
SP	São Paulo	2017	322129	699300000	4346383	12106920	1521
SP	São Paulo	2015	307705	653600000	4280837	11967825	1521
SP	São Paulo	2013	284193	582100000	4212801	11821873	1521
SP	São Paulo	2012	272493	538900000	4924895	11376685	1521
RJ	Rio de Janeiro	2017	272268	337600000	2496572	6520266	1200
RJ	Rio de Janeiro	2014	271753	300300000	2445642	6453682	1200
RJ	Rio de Janeiro	2013	266527	284300000	2424009	6429923	1197
RJ	Rio de Janeiro	2016	259890	328400000	2483787	6498837	1200
RJ	Rio de Janeiro	2012	252223	253200000	2749395	6390290	1200
DF	Brasília	2014	251938	197400000	915883	2852372	5780
RJ	Rio de Janeiro	2015	251033	320200000	2464905	6476631	1200
DF	Brasília	2017	250481	244700000	933990	3039444	5780
DF	Brasília	2015	249457	215600000	922922	2914830	5780
DF	Brasília	2013	238903	175900000	908572	2789761	5780
DF	Brasília	2016	227426	235500000	929978	2977216	5780
DF	Brasília	2012	206173	164100000	1032832	2648532	5780
PR	Curitiba	2013	90013	79767473	670649	1848946	435
PR	Curitiba	2014	88683	81198399	676033	1864416	435
PR	Curitiba	2012	85974	70637709	803583	1776761	435
PR	Curitiba	2017	79490	84702357	691568	1908359	435
PR	Curitiba	2016	77916	83746837	686612	1893997	435
RS	Porto Alegre	2013	75815	57920358	515227	1467816	497
RS	Porto Alegre	2012	73989	54204832	562121	1416714	497

Table 23: Top 25 Brightest Counties, Brazil 2012-2017

Year dropped	2012	2013	2014	2015	2016	2017	2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	NTL	NTL	NTL	NTL	NTL	NTL	NTL
Area	0.417	0.827	0.559	0.618	0.456	0.233	0.553
	(0.891)	(0.675)	(0.608)	(0.613)	(0.600)	(0.575)	(0.587)
GDP	0.311***	0.329***	0.340***	0.347***	0.431***	0.0741	0.320***
	(0.0983)	(0.0953)	(0.0963)	(0.0972)	(0.102)	(0.0927)	(0.0930)
Pop	1.920**	2.055***	1.261***	1.581***	1.564***	0.979***	1.108***
	(0.802)	(0.311)	(0.252)	(0.267)	(0.308)	(0.180)	(0.218)
GDP^2	-0.00182	-0.000580	0.000373	-0.000717	-0.00116	0.00324*	-0.00114
	(0.00196)	(0.00207)	(0.00206)	(0.00215)	(0.00225)	(0.00172)	(0.00187)
Pop^2	-0.0847***	-0.0778***	-0.0508***	-0.0596***	-0.0719***	-0.0308***	-0.0395***
	(0.0313)	(0.0115)	(0.0114)	(0.0125)	(0.0151)	(0.00867)	(0.0112)
$Area^2$	0.0355	0.0345	0.00413	0.0214	0.0130	0.0236	0.0173
	(0.0672)	(0.0514)	(0.0466)	(0.0466)	(0.0457)	(0.0443)	(0.0450)
Area×Pop	-0.0451	-0.0876***	-0.0314**	-0.0515***	-0.0260	-0.0382***	-0.0385***
	(0.0373)	(0.0221)	(0.0159)	(0.0152)	(0.0159)	(0.0132)	(0.0136)
Area ×GDP	-0.00441	-0.0163**	-0.00965	-0.0103	-0.0124	0.00202	-0.0154**
	(0.00777)	(0.00803)	(0.00823)	(0.00826)	(0.00861)	(0.00758)	(0.00771)
Pop×GDP	-0.0226***	-0.0155***	-0.0235***	-0.0219***	-0.0277***	-0.0123*	-0.0150**
	(0.00592)	(0.00472)	(0.00607)	(0.00557)	(0.00522)	(0.00737)	(0.00694)
Observations	46,474	46,468	46,468	46,468	46,468	46,468	52,038
County FE	yes	yes	yes	yes	yes	yes	yes
State×Year FE	yes	yes	yes	yes	yes	yes	yes
Number of admin areas	8,674	8,674	8,674	8,674	8,674	8,674	8,674

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 24: Placebo Test, Years Dropped