Planes, Trains, and Automobiles: Night-time Lights of the USA

Dickinson, Jeffrey

2020

Planes, Trains, and Automobiles: Night-time Lights of the USA

Jeffrey Dickinson*

Tuesday 2nd June, 2020 03:03

Abstract

This paper seeks to advance understanding of the lights-income relationship by linking the newest generation of night-time satellite images, the VIIRS images, to nationwide, panel data on 3,101 US counties, including data on both population and income. I leverage the quality and frequency of those data sources and the VIIRS lights images to decompose the links between population growth, official GDP growth, and nighttime lights growth at the county level. I use a between-county estimator to identify the effects of time-invariant infrastructure features on night-time light. Roads, rail, ports, and airports I find to be strong contributors to increases in light. I find GDP growth is weakly linked with nighttime lights though light growth is strongly linked with population growth even when controlling for substantial nonlinearities which appear to be present.

JEL Codes O51, C82, R10, R11, R12

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

*author contact: jeffrey.dickinson@graduateinstitute.ch, Department of International Economics, Graduate Institute of International and Development Studies, CP 1672 - CH-1211 Genève 1, Suisse. I am deeply grateful to my supervisors Nicolas Berman and Jean-Louis Arcand for their extensive support and comments. All remaining errors are my own.

Introduction

The literature using nighttime lights satellite images as a proxy measure for human activity dates back to the late 1970's, but the watershed papers in the economics literature were those by Henderson et al. (2012) (n=3,015), and Chen and Nordhaus (2011) (n=11,559). 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. They find a strong relationship between income and lights, this was an important finding, but it raises many important questions since this type of estimation uncovers, in my opinion, little about the true nature of the lights-income relationship. The authors in Henderson et al. (2012) also have some strong limitations on their data which is that the reference data from many poor-quality countries could be extremely noisy and make identification of the exact parameters linking income, GDP, and population difficult. Henderson et al. (2012) also notably lacked data on population growth, and did not decompose light changes into their income and population components. More recent work, using very high quality data from Sweden, has suggested that light growth is more closely linked with population movements than with fluctuations in income (Mellander et al., 2015). Levin and Zhang (2017) utilizes data from the newer VIIRS satellite, 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, but 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 has been growing. Hodler and Raschky (2014) examines 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 just come on the radar using night-time lights for localized analysis. One recently published paper 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 zone was linked with pre-existing industrial infrastructure in the surrounding area, and the existence of large markets in which to sell outputs. Lastly, 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. An overview of the capabilities and some applications of night-time lights data can be found in Donaldson and Storeygard (2016).

The principal contribution of this paper is to enhance understanding of the lights-income relationship by linking lights to panel data, including data on both population and income, of the highest quality, which are available at a fine resolution. A previous work to do that was Mellander et al. (2015) which uses aggregated firm-level data from Europe from a relatively small country, Sweden. The United States, the subject of analysis in this paper, is a much larger landmass and total population (350 million), and has substantially more heterogeneity when we consider places like Florida, Alaska, Arizona, Washington, and Hawaii. Second, I seek to enhance our understanding of what drives night-time light growth, population, and incomes, by analyzing which important infrastructure features may be driving light using a between-groups estimator, a procedure which is designed to permit identification of the effects on night-time light of geographic features which are invariant in the sample period. Although Mellander et al. (2015) mention the potential importance of electrical consumption data, the authors were not able to acquire or incorporate electrical consumption data. In this paper electrical consumption data from California are modeled as the dependent variable on an identical set of controls as the night-time lights regressions during the same sample period. Night-time light is found to be most comparable to the consumption of residential electricity, but nighttime lights data are currently available daily, for free, thus, I argue, offering substantial value-added over electrical consumption data.

Principal Hypotheses:

- What is the strength of the income-lights relationship as measured in a large diverse country with high-quality panel data?
- Does night-time light present significant value-added over electrical consumption data?
- What are the effects of certain infrastructure elements on nighttime light growth?

Methodology

Harnessing the potential of the panel structure of the income and population data is a key priority. Based on the results of other authors such as Mellander et al. (2015), night-time light appears to be a proxy for consumption more than production. The main strategy of the paper is an empirical approach to understanding 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 in the context because the satellite images from the VIIRS are somewhat noisy, while they are very precise in the dimension of how they record the texture of activity across space. As such, using the night-time lights measure as the dependent variable makes more sense, I argue, than trying to use night-time light to predict income or population size. The latter is left for future research. 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:

$$NTL = \beta_1[Income] + \beta_2[Population] + \beta_3[Area] \tag{1}$$

Based on previous papers most notably 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 the following specification, which includes squared terms and interaction terms among all three key independent variables. The third main variable besides income and population being the area of the county. The second potential specification is therefore the following:

$$NTL = \beta_1[Income] + \beta_2[Population] + \beta_3[Area] + \beta_4[Income]^2 + \beta_5[Population]^2 + \beta_6[Area]^2 \tag{2}$$

$$+\beta_7[Income] \times [Area] + \beta_8[Population] \times [Area] + \beta_9[Population] \times [Income] \tag{3}$$

Between-county Estimation

There are certain geographic characteristics of the counties which we would like to analyze, but it is difficult because these counties do not have any variation in those characteristics within the sample period. In order to obtain identification, all variables are collapsed to their 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 that lack those features. Given the size of the sample and the survey period I feel this is the most appropriate approach to consider the effects of geographic variables.

Data

Bureau of Labor Statistics GDP Data

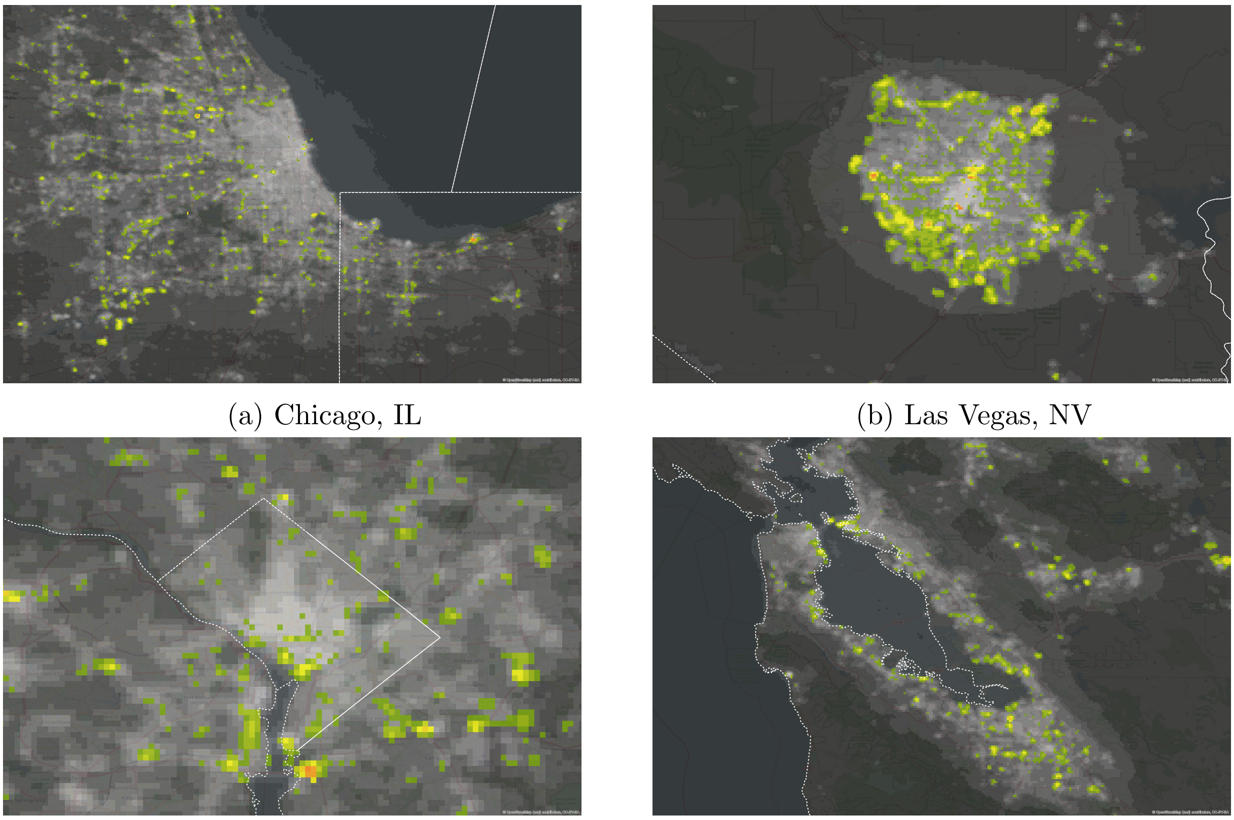
The Bureau of Labor Statistics (BLS) has recently begun releasing local-area calculations for gross domestic product. In the BLS GDP statistics, GDP is calculated using the income approach. Based on the availability of data, the Bureau of Economic Analysis (BEA) utilizes the income method for calculating 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 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 are described in Aysheshim et al. (2020).

Census Bureau ACS County-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.

LandScan Gridded Population Data

LandScan gridded population data is a global dataset in the form of an integer-based raster. The population is inferred using an algorithm and a mix of sources, the principal source being daytime satellite imagery of human settlements. These data are quite popular, and have been used in other economics research when comparable administrative population data are not available.



(a) Chicago, IL

(b) Las Vegas, NV

(c) Washington, DC

(d) San Francisco, CA

Figure 1: 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

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 great 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 2012-2019, 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 to capture imaging using the spectroradiometer, a device used to capture light similar to the capture device in a digital camera, and passes by around 1:30am and 1:30pm local time each day (Carlowicz, 2012). Raw data from the sensor are then processed to remove non-human generated disturbances such as aurora borealis, stay 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 US cities are shown in Figure 1. Long-run changes in night-time light are shown in green-red colors to demonstrate intensity. Chicago, IL is shown in the upper left panel, and is seen to be quite spread out over space. Las Vegas, NV is an interesting example because of its intensity relative to the darkness of the nearby unpopulated desert. Washington, DC provides a good illustration of how, despite high density of lights, light intensity can still be distinguished at a high resolution. The dark red spot just south of Washington, DC is National Harbor, and the major development within 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. Last, one of the wealthiest, most expensive, and most productive regions in the country is depicted in Northern California from Berkeley to San Jose, showing all the pockets of development along the way.

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.[1] 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.

Infrastructure Data

Infrastructure data, such as the location of ports, roads, 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 of Homeland Security. Airport locations were taken from open data sources.[2]

Results

The results presented in Table 1 are the principal models that have been fit to the night-time lights data. Columns 1-3 are the most parsimonious specifications, omitting all interactions. The estimated equation is the following, though in the squared terms are omitted:

$$
\begin{aligned}
NTL = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1^2 + \beta_5 x_2^2 + \beta_6 x_3^2 \\
+\beta_7 (x_1 \times x_2) + \beta_8 (x_2 \times x_3) + \beta_9 (x_1 \times x_3) + \alpha_{sy} + \epsilon_i
\end{aligned}
$$

All variables are in their log form and all columns include state·year fixed effects. Column 1 shows that area contributes significantly to the total amount of light, but so does income, almost as much as area, with population growth tracking the least strongly with night-time light. Column 2 is the county fixed-effect within estimates of the same, income and population are significant, though population has a significant and negative effect on night-time light. Column 3 is identical in model to column 2, but now using the alternative LandScan population data. We see that, at least with columns 2-3, there isn't much difference between the results for the point

[1] https://ecdms.energy.ca.gov/elecbycounty.aspx
[2] https://ourairports.com/

estimates for income with ACS population estimates and LandScan population data, though the point estimates for the effect of population on night-time light are now slightly smaller and they remain negative and statistically significant. In columns 4-6 I have now incorporated an interaction of the key variables of interest with the area variable. The intuition for this is that the relationship between lights and population or lights and income might be changing with the size of the county in question. Column 4 contains the most interesting results, possibly because the effects are not well-identified within-counties. Column 4 shows that area is increasing light significantly, the level effect of GDP is now negative and significant, population is positive and significant, though area×Population is negative and significant, area×GDP is positive and significant. It is important to note that this effect is the most important, the GDP*area effect, and this appears to be consistently estimated regardless of specification. Furthermore, the interaction on income×GDP is also significant and positive, though the effect is relatively smaller than the others.

In Table 2 we have the full specification including all interactions and squared terms. Columns 1-2 contain the ACS population data, while 3-4 contain the estimates with the LandScan population data. All of the squared terms are statistically significant in all columns. GDP appears to have increasing returns to scale within the state, but decreasing returns on the within-county estimation. With respect to the interaction terms it is more of a mixed bag. Nearly all of the interaction terms are significant. The area × GDP is positive, while area × population is negative. The effect of income × population, if present, does not appear to be identified within-county. The specifications in table 2 represent the preferred specifications, where it is not necessary to conserve degrees of freedom, given the importance of the second order terms.

Table 2 shows the results of regressions using VIIRS night-time light data as well as California electrical consumption data. 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 commercial (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.

The following table, Table 4 is the results of the between-group estimation. The procedure is used in the case that we have variables of interest which are unchanging in the period of data availability, in this case those variables are infrastructure variables. In this estimation the same fully specified model is used as before, but the ACS population variable is now dropped, since the LandScan estimates appear to perform better. These regressions are meant to demonstrate the importance of geographic variables in the production of nighttime light, which would be impossible without the use of this specialized procedure. All of the level terms are significant except for GDP, which is negative and not significant. The direct effect of population on long-term NTL appears to be very strong, 0.65-0.72. Turning to the squared terms, all of them are statistically significant and positive, indicating increasing marginal returns to all of those variables. I have also included the area*var interaction as well as population*income which means all possible interactions are included. Night-time light is decreasing in the product of area and population, meaning that as an area gets larger and as population increases, the relationship between both area and NTL and population and NTL is diminished. On the contrary, controlling for population and its derivatives we see that the affect of GDP*Area is positive and quite similar to point estimates from the previous sections, though this variable is not statistically significant at standard levels in this result. Last we have the income*population interaction term, which indicates higher population or higher income results in lower levels of night-time light. This effect appears large in the Mundlak estimates, though in the previous estimations (Table 2) it was negative and statistically significant, but the point estimate is closer to .05.

Turning to the geographic variables, these are the main variables of interest for the Mundlak regressions. The geographic control variables include an indicator variable for 1 - Ports, 2 - Principal Roads, 3 - Rail Infrastrucure Present, 4 - Airports. There are three different specifications using geographic variables which are presented. The first column contains a separate indicator for a single airport or multiple airports. It could be that most of the result was coming from counties with multiple airports, and it was for that reason that this is included. The second column includes a count of airports in the county, with more airports meaning greater night-time light. A marginal airport increases like by .0286, statistically significant at the .01 level. Lastly, I include an interaction variable that indicates if a county has all of the following: a port, a major road, rail infrastructure, and at least one airport. This variable then indicates that the simultaneous presence of all of those infrastructure elements are a determinant of night-time light. The point estimates of this effects are .0556, statistically significant at the .01 level. This result can be thought of as the additional marginal benefit of having the combined presence of ports, roads, rail, and airports, over having an individual airport, port, rail, or road.

Conclusion

Using administrative and survey data of the highest quality, 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. I argue night-time light is found to be a strong proxy indicator for population changes, while it is only a weak indicator for changes in income. Although night-time light only moves slightly with income, light still moves more consistently with income than does electrical consumption data, and night-time light may therefore still be a useful proxy indicator for changes in population or income in small geographic areas for which accurate and timely data are not available. Infrastructure elements are found to substantially influence light production, a finding which could be useful to future researchers looking to use VIIRS imagery for economic analysis.

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).

1 Appendix

County-Level Nighttime Lights Regressions

VARIABLES	N	mean	sd	min	max	skewness	kurtosis	p5	p25	p50	p75	p95
sum_nl	21,707	17,458	67,003	447	2922000	23.1	740.4	1,779	3,589	6,469	13,451	57,182
bls_gdp	21,707	5538000	24260000	2,753	710900000	15.18	327.1	110,712	336,120	877,472	2620000	22140000
sum_pop	21,707	103,141	333,895	81	10140000	13.55	308.7	2,608	10,603	24,987	67,842	436,437
acs_pop	21,707	104,286	332,582	86	10120000	13.52	309.4	2,879	11,028	25,999	68,919	442,728
distance_road	21,707	0.381	2.469	0	39.83	12.39	162.6	0	0	0.0295	0.255	0.835
f500_hq	21,707	0.05	0.218	0	1	4.13	18.06	0	0	0	0	0
airport_count	21,707	0.479	1.19	0	27	9.152	147	0	0	0	1	2
square_feet	21,707	2985000000	9591000000	40570000	380900000000	26.07	902.7	605700000	1149000000	1646000000	2458000000	831300000
square_miles	21,707	1,152	3,703	15.67	147066	26.07	902.7	233.9	443.6	635.7	949.1	3,210
log_bls_gdp	21,707	13.86	1.604	7.921	20.38	0.61	3.428	11.61	12.73	13.68	14.78	16.91
log_sum_pop	21,707	10.25	1.505	4.407	16.13	0.303	3.328	7.867	9.269	10.13	11.12	12.99
log_sum_nl	21,707	8.932	1.071	6.105	14.89	0.86	4.238	7.484	8.186	8.775	9.507	10.95
log_area	21,707	6.555	0.797	2.813	12	0.951	6.824	5.459	6.097	6.456	6.857	8.074
log_acs_pop	21,707	10.29	1.486	4.466	16	0.308	3.34	7.966	9.308	10.17	11.14	13
has_port	21,707	0.0271	0.162	0	1	5.826	34.94	0	0	0	0	0
has_rail	21,707	0.881	0.324	0	1	-2.353	6.539	0	1	1	1	1
has_road	21,707	0.45	0.498	0	1	0.199	1.04	0	0	0	1	1
has_airport	21,707	0.315	0.465	0	1	0.795	1.631	0	0	0	1	1
has_multiple_airports	21,707	0.0713	0.257	0	1	3.333	12.11	0	0	0	0	1
has_navigable_water	21,707	0.304	0.46	0	1	0.852	1.725	0	0	0	1	1
has_all_four	21,707	0.078	0.268	0	1	3.146	10.9	0	0	0	0	1

Table 1: Variables Used in County-Level Night-time Lights Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	log_sum_nl	log_sum_nl	log_sum_nl	log_sum_nl	log_sum_nl	log_sum_nl
Area	0.444***			1.234***		
	(0.00757)			(0.0474)		
BLS GDP	0.475***	0.0717***	0.0688***	-0.534***	-0.0193	-0.0664
	(0.00889)	(0.0275)	(0.0266)	(0.0580)	(0.243)	(0.241)
ACS Pop	0.153***	-0.480***		0.821***	-1.370**	
	(0.0105)	(0.104)		(0.0674)	(0.660)	
LS Pop			-0.360***			0.387
			(0.0603)			(0.335)
Area × ACS Pop				-0.178***	0.208**	
				(0.00911)	(0.1000)	
Area × BLS GDP				0.0775***	0.0522	0.0548*
				(0.00847)	(0.0333)	(0.0329)
ACS Pop × BLS GDP				0.0407***	-0.0334***	
				(0.00127)	(0.00877)	
Area × LS Pop						-0.0486
						(0.0486)
LS Pop × BLS GDP						-0.0300***
						(0.00888)
Observations	21,707	21,707	21,707	21,707	21,707	21,707
R-squared	0.901	0.685	0.688	0.927	0.690	0.692
Number of Counties		3,101	3,101		3,101	3,101
State-Year FE	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses

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

Table 2: County-Level Night-time Lights Regressions

VARIABLES	(1) Total NTL	(2) Total NTL	(3) Total NTL	(4) Total NTL
Area	0.410***		0.357***	
	(0.0547)		(0.0583)	
BLS GDP	0.0340	0.219	0.0161	0.281
	(0.0822)	(0.268)	(0.0903)	(0.274)
ACS Pop	0.277***	2.233***		
	(0.0884)	(0.679)		
LS Pop			0.345***	0.952**
			(0.0937)	(0.415)
BLS GDP^2	0.0147**	-0.0309***	0.0195***	-0.0265**
	(0.00714)	(0.00886)	(0.00577)	(0.0103)
ACS Pop^2	0.0721***	-0.190***		
	(0.00719)	(0.0316)		
LS Pop^2			0.0942***	-0.0413**
			(0.00507)	(0.0180)
Area^2	0.0613***		0.0612***	
	(0.00359)		(0.00372)	
Area × BLS GDP	0.0461***	0.0741**	0.0609***	0.0662**
	(0.0108)	(0.0331)	(0.0112)	(0.0337)
Area × ACS Pop	-0.136***	0.160**		
	(0.0109)	(0.0754)		
Area × LS Pop			-0.152***	-0.0792
			(0.0110)	(0.0485)
ACS Pop × BLS GDP	-0.0408***	0.0159		
	(0.0127)	(0.0178)		
LS Pop × BLS GDP			-0.0667***	0.00236
			(0.00892)	(0.0218)
Observations	21,707	21,707	21,707	21,707
R-squared	0.934	0.696	0.935	0.695
Number of Counties		3,101		3,101
State-Year FE	yes	yes	yes	yes

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

Table 3: County-Level Night-time Lights Regressions, Full Specification

VARIABLES	N	mean	sd	min	max	skewness	kurtosis	p5	p25	p50	p75	p95
elec_cons1	406	3,315	7,021	4.008	49,193	4.77	29.85	43.95	240.4	781.4	3,200	13,026
log_elec_cons1	406	6.712	1.833	1.388	10.8	-0.226	2.95	3.783	5.482	6.661	8.071	9.475
elec_cons2	406	1,585	3,090	9.291	21,162	4.438	26.43	15.67	200	553.2	1,512	6,859
log_elec_cons2	406	6.223	1.629	2.229	9.96	-0.233	2.902	2.751	5.298	6.316	7.321	8.833
elec_cons3	406	4,901	10,032	13.89	69,946	4.728	29.54	113.1	410.2	1,474	4,763	19,627
log_elec_cons3	406	7.231	1.71	2.631	11.16	-0.151	2.816	4.728	6.017	7.295	8.469	9.885
sum_nl	406	54,822	112,144	755.6	822,111	4.562	27.4	2,834	6,380	17,507	48,619	238,090
bls_gdp	406	41730000	97600000	47224	710900000	4.613	27.66	426486	1558000	7615000	27860000	210700000
sum_pop	406	668138	1453000	1140	10140000	4.893	30.72	10068	46025	181767	713034	3108000
acs_pop	406	669915	1452000	1057	10120000	4.896	30.77	9355	44957	181536	721929	3080000
square_feet	406	7063000000	8020000000	125800000	52100000000	3.422	18.22	1169000000	2529000000	4024000000	9116000000	21120000000
square_miles	406	2,727	3,097	48.56	20,118	3.422	18.22	451.5	976.6	1,554	3,520	8,154
square_km	406	7,063	8,020	125.8	52,104	3.422	18.22	1,169	2,529	4,024	9,116	21,118
log_bls_gdp	406	15.86	1.983	10.76	20.38	0.00906	2.434	12.96	14.26	15.85	17.14	19.17
log_count	406	10.21	0.956	6.616	12.6	-0.499	5.061	8.833	9.63	10.08	10.91	11.7
log_sum_pop	406	12.03	1.824	7.04	16.13	-0.245	2.83	9.217	10.74	12.11	13.48	14.95
log_sum_nl	406	9.862	1.4	6.629	13.62	0.38	2.653	7.95	8.761	9.77	10.79	12.38
log_area	406	7.479	0.959	3.903	9.909	-0.462	4.999	6.115	6.885	7.349	8.166	9.006
log_acs_pop	406	12.04	1.828	6.964	16.13	-0.276	2.894	9.144	10.71	12.11	13.49	14.94

Table 4: Summary Statistics of Variables Used in Electrical Consumption Regressions

VARIABLES	(1) Log Total NTL	(2) Log Total NTL	(3) Log Total Elec	(4) Log Total Elec.	(5) Log Resid. Elec.	(6) Log Resid. Elec.	(7) Log Comm. Elec.	(8) Log 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	0.604	0.981	0.075	0.956	0.042	0.964	0.275
Number of FIPS		58		58		58		58

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

Table 5: California Electrical Consumption Regressions

VARIABLES	(1) Total NTL	(2) Total NTL	(3) Total NTL
Area	0.357**	0.481***	0.341**
	(0.172)	(0.165)	(0.173)
BLS GDP	-0.305	-0.318	-0.294
	(0.269)	(0.265)	(0.270)
LS Pop	0.652**	0.720***	0.657**
	(0.263)	(0.258)	(0.264)
BLS GDP^2	0.0450***	0.0461***	0.0446***
	(0.0155)	(0.0153)	(0.0154)
LS Pop^2	0.103***	0.103***	0.103***
	(0.00865)	(0.00874)	(0.00864)
Area^2	0.0616***	0.0537***	0.0620***
	(0.00950)	(0.00908)	(0.00946)
Area $\times$ LS Pop	-0.152***	-0.156***	-0.152***
	(0.0316)	(0.0312)	(0.0317)
Area $\times$ BLS GDP	0.0612*	0.0625*	0.0617*
	(0.0328)	(0.0320)	(0.0329)
LS Pop $\times$ BLS GDP	-0.104***	-0.107***	-0.104***
	(0.0179)	(0.0178)	(0.0179)
Has Port	0.123***	0.122***	0.110***
	(0.0305)	(0.0298)	(0.0304)
Has Road	0.105***	0.106***	0.101***
	(0.0102)	(0.0102)	(0.0102)
Has Rail	0.0624***	0.0629***	0.0622***
	(0.0174)	(0.0173)	(0.0174)
F500 HQ	-0.0473	-0.0507	-0.0449
	(0.0368)	(0.0358)	(0.0366)
Has Airport	-0.00172		-0.0106
	(0.0130)		(0.0133)
Has Multiple Airports	0.0181		0.0139
	(0.0251)		(0.0250)
Airport Count		0.0286***	
		(0.00957)	
Has All Four			0.0556***
			(0.0209)
Observations	3,101	3,101	3,101
R-squared	0.950	0.950	0.950
State-FE	yes	yes	yes

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

Table 6: Mundlak Procedure: Economic Geography Variables

VARIABLES	(1) log_sum_nl	(2) log_sum_nl	(3) log_sum_nl
Alabama	-	-	-
Alaska	0.583***	0.480***	0.594***
	(0.161)	(0.164)	(0.162)
Arizona	-0.666***	-0.633***	-0.657***
	(0.0808)	(0.0731)	(0.0806)
Arkansas	-0.0716**	-0.0720**	-0.0715**
	(0.0306)	(0.0304)	(0.0306)
Californa	-0.940***	-0.933***	-0.938***
	(0.0490)	(0.0484)	(0.0493)
Colorado	-0.477***	-0.468***	-0.470***
	(0.0413)	(0.0414)	(0.0415)
Connecticut	-0.811***	-0.805***	-0.819***
	(0.0585)	(0.0556)	(0.0593)
Delaware	-0.359***	-0.357***	-0.373***
	(0.0995)	(0.0953)	(0.111)
District of Columbia	-0.659***	-0.601***	-0.629***
	(0.230)	(0.229)	(0.232)
Florida	-0.287***	-0.294***	-0.297***
	(0.0299)	(0.0303)	(0.0302)
Georgia	-0.0386	-0.0422	-0.0358
	(0.0292)	(0.0285)	(0.0294)
Hawaii	-1.425***	-1.493***	-1.399***
	(0.219)	(0.213)	(0.218)
Idaho	-0.471***	-0.461***	-0.468***
	(0.0415)	(0.0416)	(0.0419)
Illinois	-0.117***	-0.114***	-0.116***
	(0.0284)	(0.0281)	(0.0285)
Indiana	-0.145***	-0.145***	-0.141***
	(0.0283)	(0.0279)	(0.0288)
Iowa	-0.169***	-0.168***	-0.167***
	(0.0254)	(0.0251)	(0.0256)
Kansas	-0.293***	-0.297***	-0.288***
	(0.0280)	(0.0274)	(0.0284)
Kentucky	-0.0879***	-0.0868***	-0.0883***
	(0.0285)	(0.0282)	(0.0286)
Louisiana	0.0955**	0.101**	0.0947**
	(0.0426)	(0.0426)	(0.0427)
Maine	-0.798***	-0.790***	-0.794***
	(0.0488)	(0.0505)	(0.0490)
Maryland	-0.461***	-0.458***	-0.455***
	(0.0395)	(0.0396)	(0.0395)
Massachusetts	-0.901***	-0.892***	-0.908***
	(0.0844)	(0.0827)	(0.0822)
Michigan	-0.336***	-0.337***	-0.333***
	(0.0334)	(0.0332)	(0.0337)
Minnesota	-0.108***	-0.102***	-0.106***
	(0.0298)	(0.0295)	(0.0301)
Observations	3,101	3,101	3,101
R-squared	0.950	0.950	0.950
State-FE	yes	yes	yes

Robust standard errors in parentheses

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

Table 4: Mundlak Procedure - State Fixed Effects Estimates

VARIABLES	(1) log_sum_nl	(2) log_sum_nl	(3) log_sum_nl
Mississippi	0.0586**	0.0571**	0.0606**
	(0.0279)	(0.0274)	(0.0281)
Missouri	-0.166***	-0.165***	-0.165***
	(0.0246)	(0.0243)	(0.0248)
Montana	-0.276***	-0.263***	-0.269***
	(0.0472)	(0.0457)	(0.0476)
Nebraska	-0.208***	-0.207***	-0.205***
	(0.0314)	(0.0307)	(0.0315)
Nevada	-0.644***	-0.611***	-0.636***
	(0.0975)	(0.0941)	(0.0978)
New Hampshire	-0.869***	-0.866***	-0.863***
	(0.0645)	(0.0653)	(0.0627)
New Jersey	-0.570***	-0.555***	-0.565***
	(0.0700)	(0.0698)	(0.0706)
New Mexico	-0.330***	-0.307***	-0.322***
	(0.0649)	(0.0631)	(0.0652)
New York	-0.729***	-0.719***	-0.734***
	(0.0468)	(0.0453)	(0.0469)
North Carolina	-0.244***	-0.245***	-0.239***
	(0.0307)	(0.0304)	(0.0311)
North Dakota	0.169**	0.174**	0.173**
	(0.0786)	(0.0785)	(0.0787)
Ohio	-0.174***	-0.172***	-0.171***
	(0.0281)	(0.0279)	(0.0284)
Oklahoma	-0.00667	-0.00694	-0.00427
	(0.0290)	(0.0288)	(0.0293)
Oregon	-0.843***	-0.827***	-0.841***
	(0.0519)	(0.0510)	(0.0522)
Pennsylvania	-0.566***	-0.561***	-0.562***
	(0.0320)	(0.0320)	(0.0321)
Rhode Island	-0.670***	-0.665***	-0.690***
	(0.0806)	(0.0843)	(0.0789)
South Carolina	0.0295	0.0261	0.0295
	(0.0361)	(0.0364)	(0.0360)
South Dakota	-0.134***	-0.131***	-0.130***
	(0.0297)	(0.0293)	(0.0301)
Tennessee	-0.0926***	-0.0900***	-0.0935***
	(0.0282)	(0.0278)	(0.0283)
Texas	-0.0704**	-0.0675**	-0.0653**
	(0.0281)	(0.0279)	(0.0284)
Utah	-0.490***	-0.472***	-0.482***
	(0.0588)	(0.0579)	(0.0590)
Vermont	-0.944***	-0.943***	-0.942***
	(0.0840)	(0.0840)	(0.0842)
Virginia	-0.410***	-0.404***	-0.410***
	(0.0372)	(0.0370)	(0.0373)
Washington	-0.794***	-0.784***	-0.798***
	(0.0467)	(0.0474)	(0.0468)
West Virginia	-0.307***	-0.309***	-0.308***
	(0.0398)	(0.0399)	(0.0397)
Wisconsin	-0.206***	-0.207***	-0.203***
	(0.0268)	(0.0265)	(0.0272)
Wyoming	-0.465***	-0.441***	-0.458***
	(0.0638)	(0.0616)	(0.0635)
Observations	3,101	3,101	3,101
R-squared	0.950	0.950	0.950
State-FE	yes	yes	yes

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

Table 4: Mundlak Procedure - State Fixed Effects Estimates Ctd.

State	County	year	count	Mean NL	Total NL	BLS GDP	Pop	sum_pop	nl_percap	gdp_
District of Columbia	District of Columbia	2014	994	78.95	78476		3402.52	843824	0.09	
District of Columbia	District of Columbia	2013	994	73.29	72852		3357.11	832563	0.09	
District of Columbia	District of Columbia	2015	994	72.37	71939	117200000	3467.83	860021	0.08	13
District of Columbia	District of Columbia	2017	994	70.21	69787	121000000	3700.80	917799	0.08	13
District of Columbia	District of Columbia	2019	994	69.66	69245					
New York	New York	2014	421	67.67	28487		22959.81	2387820	0.01	
New York	New York	2015	421	65.99	27784	573500000	23070.30	2399311	0.01	23
New York	New York	2017	421	65.84	27719	601500000	24194.27	2516204	0.01	23
New York	New York	2013	421	64.76	27265		22866.47	2378113	0.01	
New York	New York	2019	421	63.01	26525					
District of Columbia	District of Columbia	2018	994	59.89	59528	124000000	3738.09	927047	0.06	13
New York	New York	2016	421	59.43	25019	585700000	23835.62	2478904	0.01	23
New York	New York	2018	421	59.06	24863	600200000	24082.68	2504599	0.01	23
District of Columbia	District of Columbia	2016	994	58.16	57814	119600000	3610.92	895509	0.06	13
New Jersey	Hudson	2014	774	56.23	43520		3155.95	634345	0.07	
New Jersey	Hudson	2015	774	54.19	41947		3181.29	639440	0.07	
New Jersey	Hudson	2019	774	52.33	40505					
New Jersey	Hudson	2017	774	50.04	38728		3253.12	653877	0.06	
Pennsylvania	Philadelphia	2014	2242	49.59	111187		2896.39	1621978	0.07	
New York	Bronx	2014	715	49.05	35072		7262.42	1285449	0.03	
New Jersey	Hudson	2013	774	47.99	37143		3113.44	625801	0.06	
Illinois	Cook	2014	15567	47.79	743964		1384.09	5403468	0.14	
Virginia	Norfolk	2014	853	47.72	40704		1178.46	251011	0.16	
New York	Bronx	2015	715	47.26	33794	38864519	7332.82	1297909	0.03	29

Table 5: Top 25 US Counties in Mean Light per Pixel 2013-2019

name_1	name_2	year	count	mean_nl	sum_nl	bls_gdp	mean_pop	sum_pop	nl_percap	gdp_percap
Hawaii	Kalawao	2016	212	0.03	6.12		1	67	0.09	
Nevada	Esmeralda	2016	54708	0.05	2615.11		0	693	3.77	
California	Trinity	2016	50953	0.05	2456.68	473471	1	11974	0.21	39.54
Texas	Jeff Davis	2016	32122	0.05	1709.73	70060	0	2538	0.67	27.60
New Mexico	Catron	2016	101015	0.05	5404.51	74135	0	3260	1.66	22.74
Nevada	Lincoln	2016	162114	0.05	8887.32	168014	0	5331	1.67	31.52
New Mexico	Harding	2016	31678	0.06	1746.63		0	650	2.69	
Michigan	Lake Superior	2016	284016	0.06	15773.81		24	10165	1.55	
Oregon	Lake	2016	136772	0.06	7667.22	284328	0	7724	0.99	36.81
Oregon	Wheeler	2016	28996	0.06	1631.34	53876	0	1269	1.29	42.46
Texas	Brewster	2016	86232	0.06	4913.46	350778	0	10286	0.48	34.10
Utah	Kane	2016	62287	0.06	3590.75	266590	0	7354	0.49	36.25
Oregon	Harney	2016	168877	0.06	9767.36	254129	0	6957	1.40	36.53
New Mexico	Debaca	2016	34169	0.06	2008.12		0	1539	1.30	
Texas	Edwards	2016	29434	0.06	1730.72	106239	0	1753	0.99	60.60
California	Inyo	2016	153717	0.06	9045.35	1184726	0	18205	0.50	65.08
Texas	Presidio	2016	53658	0.06	3223.45	254849	1	7085	0.45	35.97
Utah	Garfield	2016	79714	0.06	4882.74	206823	0	5215	0.94	39.66
California	Modoc	2016	67674	0.06	4176.31	445488	1	8523	0.49	52.27
Maine	Piscataquis	2016	75601	0.06	4765.02	495156	1	16508	0.29	29.99
Utah	Piute	2016	11792	0.06	743.64	43413	1	1481	0.50	29.31
Nevada	Mineral	2016	58769	0.06	3719.57	209668	0	4228	0.88	49.59
Utah	Wayne	2016	37760	0.06	2415.96	90479	0	2695	0.90	33.57
New Mexico	Mora	2016	28854	0.06	1861.48	105835	1	3725	0.50	28.41
Nebraska	Arthur	2016	11671	0.07	764.48		0	414	1.85	

Table 6: Bottom 25 US Counties in Mean Light per Pixel 2013-2019