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]American University

Abstract

The newest generation of nighttime lights satellite images offer a resolution 45 times higher than the previous generation. This paper links those images to nationwide panel data on population and income from the United States at the county level, and Brazil at the município level, for the years 2012-2020. Controlling for the direct effect of population on light, I confirm that nighttime light responds strongly to changes in income at a high resolution. Importantly, in Brazil, except for the highest output areas, the effect of changes in local *population* track more strongly with nighttime lights than do changes in local economic output. I use a between-county estimator to provide identification of the effects of time-invariant infrastructure features on night-time light. My estimates suggest that railways are associated with lower levels of nighttime light while border crossings contribute positively and significantly to nighttime light.
JEL Codes O1, O18, R12
Keywords: night-time light, GDP, population, infrastructure, regional development, Brazil, USA

*Correspondence: dickinso@american.edu, Department of Economics, American University, 4400 Massachusetts Ave., NW Washington, D.C. 20016. Much of the work that contributed to this paper was completed during graduate studies at the Graduate Institute of International and Development Studies in Geneva, Switzerland. None of this work would have been possible without help from my doctoral supervisors, Jean-Louis Arcand, who introduced me to the subject of nighttime lights and impact evaluation, and Nicolas Berman, an incredible applied economist, a patient teacher and a supportive mentor. I am grateful for their extensive support, academic and otherwise, and many, many thoughtful comments. I also have benefited from feedback from Michele Andreolli. All remaining errors are my own.

1 Introduction

Catalyzed by groundbreaking papers from Henderson et al. (2012) and Chen and Nordhaus (2011), over the past decade, the use of data on human activity extracted from nighttime satellite images has flourished among social scientists. Nighttime lights images are free, high-quality, widely available and capable of providing insights that other data are unable to provide. Most significantly, nighttime lights provide high-resolution data on changes in human activity for most of the globe. Social scientists have already been utilizing nighttime lights as a proxy for GDP at a high resolution for some time (Hodler and Raschky, 2014; Kocornik-Mina et al., 2020). To the best of my knowledge, however, the strength of the GDP-lights relationship has never been tested using panel data at the sub-national level. Performing panel data analysis of the lights-GDP relationship is meaningful, I argue, because it provides support for the many researchers interested in taking advantage of nighttime lights data at the sub-national level. As an example of where the strength of the lights-GDP relationship may break down, authors Huang et al. (2021) analyze the effect of a large, randomized, unconditional cash transfer program on nighttime lights using the highest-resolution nighttime lights data. Despite the randomized design of the intervention, the authors find that nighttime light is a weak proxy in rural areas of Kenya. In those areas, where electrical infrastructure is poor, it could be nighttime light is more related to the size of the local population than it is with changes in consumption. For this reason, it is important to understand with greater clarity the link between changes in population, changes in GDP, and changes in nighttime light, so that researchers can apply these data when it suits their analysis. This paper, therefore, aims to reconcile the issue of whether nighttime light is indeed a valid and strong proxy for GDP in small areas such as U.S. counties and Brazilian municípios.

Though nighttime lights data are seen as a tool to measure countries which lack high-quality administrative data, a lack of high-quality administrative data also makes it difficult to validate that nighttime light responds to changes in economic output. Authors in the nighttime lights literature have analyzed the strength of the lights/GDP relationship thus far by stratifying estimates across different population densities or using population density at a particular point in time. The literature has thus not yet estimated the *direct* effect of population on light using panel data, prior authors have only examined the effect of population indirectly or with cross-sectional estimates. If changes in population relate more closely to changes in night light than do changes in GDP, it is important to directly incorporate population as a control variable in the econometric model to make inference about the direct GDP-lights relationship. I argue this type of validation is critical to carry out in order to precisely estimate of the elasticity of nighttime lights with respect to GDP.

If GDP and nighttime lights are more strongly related in one country than another, that has implications for nighttime lights estimation at the trans-national (country) level. If the relationship is indeed stronger in the U.S.A. than in Brazil, an increase in light in a U.S. county has a different meaning than a corresponding increase in light in a Brazilian município. By comparing NTL-GDP dynamics in both Brazil and the U.S.A., I assess the differences in the observed relationships between these variables. Importantly, if the effect of population on nighttime light is stronger than the effect of GDP on nighttime light, researchers may be proxying for something entirely different than that which was intended. Even with powerful fixed-effects procedures, combining countries in nighttime lights regressions may be problematic if nighttime lights measure population better in some places GDP better in others.

To bring clarity to the question of the sub-national nighttime lights GDP relationship, I apply panel data fixed-effects techniques and incorporate spatial econometrics as well as higher-order terms. To the best of my knowledge, panel data and spatial econometrics techniques have not yet been applied to sub-national nighttime lights data to test the GDP-lights relationship. The omission of multi-period fixed effects in previous estimates with VIIRS data leaves prior estimates vulnerable to bias from unobserved variables such as public goods, terrain, or changes in the level of prices in terms of electricity or overall inflation. My results confirm that nighttime light *is* a valid proxy for GDP at a high spatial resolution for the U.S.A. and for some parts of Brazil. An important caveat of my results is that, in Brazil, where the informal sector is large and electrical access is limited, I estimate changes in the size of the *population* have a stronger relationship with nighttime light than changes in município-level GDP.

Though their use has not yet become widespread, the newest generation of images, known as VIIRS nighttime lights, offer consequential advancements over the previous generation of images. Unlike it's predecessor, the latest light-capturing sensor was purpose-built for capturing nighttime images of human activity.[1] Improvements to the sensor include greatly increased sensitivity at both the extensive and intensive margins of light. This amplification of sensor accuracy is of importance to researchers and analysts wishing to proxy for GDP in a small area (Donaldson and Storeygard, 2016; Gibson et al., 2021). Data from the previous generation of satellites were limited such that the sensor was unable to record values beyond a certain threshold. This resulted in dense and bright areas, particularly urban areas, not being as precisely measured.[2] VIIRS lights images no longer face this limitation, which represents a major advantage for researchers interested in analyzing changes within urban areas (Shi et al., 2014; Chen and Nordhaus, 2015; Elvidge et al., 2017).[3]

Authors in the economics and remote sensing research communities have, in recent years, scrutinized the relationship between GDP and nighttime lights. Previous nighttime lights papers have primarily assessed the correspondence of GDP and light without controlling directly for the effect of changes in population on nighttime light. For example, Chen and Nordhaus (2019) find that the relationship between VIIRS light and GDP is strong at the metropolitan level using USA data, while Gibson and

[1]Previous satellite images were captured on-board the satellites as part of the Defense Meteorological Satellite Program (DMSP). The satellite's original intended purpose was collecting images of clouds at night for tracking weather systems.

[2]Example images can be found in appendix figure 3

[3]VIIRS stands for Visible Infrared Imaging Radiometry Suite

Boe-Gibson (2021) find that the relationship between GDP and nighttime light is fairly weak at the county level in the USA using a combination of older satellite data and VIIRS data. In some cases, policymakers and economists analyze aspects of the economy beyond gross output. Many analysts would like to make inference about changes in relative welfare or output per worker. How much of the change in light is driven by migration, for example, and how much by increases in consumption or output per worker? Without an econometric model that incorporates both elements, we cannot know the answer. Probing the validity of nighttime lights as a measure for GDP while controlling for population is important in the context of remote places such as a village in Africa undertaking a project to install a generator, such as in Huang et al. (2021). If nighttime light is capturing exclusively or largely changes in population, this might indicate nighttime lights data is of little value in measuring such a project. By decomposing these relationships at a high geospatial resolution, and accounting for the explicit spatial relationships that are present in the data, future users of nighttime light will have a sharper understanding of whether nighttime light is capturing changes in economic output or changes in the number of people.

To advance the literature on the relationship between nighttime light and GDP, I believe this paper makes three key contributions. Using data from the USA and Brazil, I first analyze the relative contribution to nighttime light of population changes and contemporaneous changes in GDP using within-county and within-município estimates. County and município data are high-resolution, and this result lends support for any researchers who wish to conduct difference-in-differences analysis or other analysis with nighttime lights at a local level. Using data with this level of precision permits me to identify whether VIIRS nighttime lights changes relate more closely to changes in the population of an area, or changes in the economic output at a high geospatial resolution.[4]

[4]Counties are the second administrative unit in the United States while municípios, literally 'municipalities,' are their counterpart from Brazil.

As an additional step, I employ several econometric strategies new to nighttime lights - GDP estimates. As I highlighted previously, departing from the previous literature, I estimate directly the effect of population on nighttime light in the econometric model. Importantly, this accounts for omitted variable bias due to the omission of population data. I apply non-parametric standard error adjustments based on other applied microeconomics papers such as Berman et al. (2017) and Hsiang (2010), to account for cross-sectional spatial inter-dependencies among counties and municípios. The Conley standard error procedure allows for location-specific serial correlation (Conley, 1999). Additionally, I utilize state×year fixed-effects that control for *time-variant* state-year specific heterogeneity such as price shocks, weather shocks or state-wide political elections or policies. Due to the findings of Hu and Yao (2021) I account for nonlinear relationships between GDP, population and nighttime light. I include higher-order transformations of the control variables as part of my estimates and present the results alongside a parsimonious model where relevant.

A final contribution is drawing a contrast between nighttime lights estimates from the USA and Brazil, two countries with some degree of overlap in physical and economic characteristics, and both countries which play a significant role in the global economy. The average município size is about 1.5km^2, while the average size of a U.S.A. county is 3km^2, though there is overlap in these categories as demonstrated by the comparisons in table 3. By comparing municípios and counties of similar size in the USA and Brazil, we can see if the relationships between population, GDP and light are consistent for the two countries, or if, in the case of one country, nighttime light might measure population changes better while in the other it better captures changes in economic output.

An essential element of this paper, I believe, is therefore to directly include within-county or within-município changes in GDP *and* within-county or within-município changes in population size in the estimated model. In other words, no other paper has used nighttime lights panel data to disaggregate the partial relationship between population and high-resolution nighttime light (Mellander et al., 2015; Gibson and Boe-Gibson, 2021). This is significant because regressions of nighttime light on GDP or GDP on nighttime light which omit variables such as population that influence, *a priori*, both GDP and nighttime light may suffer from endogeneity bias. This classical omitted variable bias implies the potential for a breakdown of the Gauss-Markov assumptions, as the error term of a regression omitting population would then be known to be correlated with GDP *and* with light via changes in population.

To reiterate, departing from previous literature on the nighttime lights-GDP relationship, in this paper I apply popular spatial econometrics techniques to provide the most precise possible estimates of the elasticity of nighttime light with respect to GDP at a high geospatial resolution. I include controls for higher-order transformations of the control variables, including the interaction of GDP and population variables. The depth of this analysis is supported by the size and characteristics of these datasets. In a spatially dense area it is likely that county and município-level economic shocks are correlated across space as well as time. If there are inter-dependencies between a unit's unobservable characteristics and our variables of interest in an estimation procedure, this can affect the parameter and standard error estimates (Conley, 1999). Estimation procedures which do not account for the spatial structure of units of observation omit important aspects of economic relationships and may suffer from a lack of precision. To that end, I employ a procedure developed by Conley (1999) which utilizes a non-parametric Generalized Method of Moments (GMM) technique to account for spatial dependence among counties and municípios. This technique permits the modeling of complex inter-relationships across individuals within the dataset. The Conley standard error technique has been leveraged in other applied economics work such as Hsiang (2010), Berman et al. (2017) and Egger et al. (2019). I take advantage of the flexibility this estimator offers to allow for the spatial effects of economic shocks up to 5,500km (3,417 miles) from the unit of observation (county or município). This is a very large distance, though it makes sense as both countries are large and economic shocks could be well-integrated. Within the framework of the spatial error model this implies economic shocks in Alaska can influence the western half of the United States, for example, but not the mid-Atlantic states like Maryland, Pennsylvania, New York, New Jersey and Delaware. The Conley standard-error estimation procedure used here also allows for location-specific serial correlation meaning shocks whose effects dissipate over several periods rather than after a single period.

To obtain the most accurate estimates it makes sense to leverage the panel dimension of the data in a panel data fixed-effects model. The within-county or within-município estimator is a meaningful tool in this context as it controls for time-invariant, unobserved, individual heterogeneity in counties and municípios. This includes elevation, terrain, and the presence of infrastructure or public goods. As long as the composition of the county or município's GDP is relatively stable over the sample period, 2012-2020, county and and município-level fixed-effects account for differences in the composition of county and município economic activity. This econometric setup allows a direct estimate of the effect of GDP and population on nighttime light. County and município-level fixed-effects also control for other time-invariant, unobserved, county-specific characteristics such as the level of human capital, other features including the presence of the state capital or steep changes in elevation. Using the fixed-effects within procedure also, in theory, can account for measurement error in GDP and population estimates. If the error remains constant over the sample period, it will be controlled for by the inclusion of fixed-effects.

The size of the data in my sample and the extensive coverage of the VIIRS nighttime lights over 8 annual images supports enhanced fixed-effects specifications: there are many observations per state and we observe many state-years over time. It is therefore possible to control for state-year unobserved heterogeneity in addition to the controls for county-and-município-level unobserved heterogeneity. Most importantly, state$\times$year fixed effects account for *time-variant*, state-year specific heterogeneity that would not otherwise be captured in a normal fixed-effects within-county transformation.[5] The inclusion of state$\times$year dummies therefore accounts for state-specific annual shocks such as weather shocks or regional political shocks (elections). To the best of my knowledge, this is the first VIIRS nighttime lights-GDP analysis that includes state$\times$year fixed effects combined with within-county and within-município transformations.[6]

A third contribution of this paper is to compare the United States nighttime lights estimates side-by-side with contemporaneous estimates from quality data provided to the public by the Brazilian statistical agency. In analyzing both the USA and Brazil I look to broaden the analysis of nighttime lights outside of the developed world by focusing on a large middle-income economy with a greater amount of poverty than the USA. By examining the strength of the relationships separately in the USA and Brazil I can draw contrast between the relative strength of the relationships between GDP, population and nighttime light in the two countries. An interesting point of contrast is that the Brazilian economy has a much larger share of informal sector firms and laborers. In 2019 there were estimated to be 38.4 million workers in Brazil's informal sector or about 41.1% of workers versus 6.3% of workers in the USA (Elgin and Yu., 2021).[7] In theory this should lead to a weaker relationship between GDP and nighttime light in Brazil as a larger share of economic activity will not be measured in GDP statistics. This is exactly what I observe in the data.

I chose these countries because the U.S. Bureau of Economic Analysis (BEA) provides high-quality, high-resolution annual population estimates and GDP data, as does the Brazilian statistical authority, the Instituto Brasileiro de Geografia e Estatística. Between them the two countries boast a combined population of around 500 million persons, covering 6.7% of the world's population. Brazilian municípios and American

[5]An example of the benefit of this is that previous work has raised the need to include the price of electricity in estimates of the elasticity of nighttime light with respect to GDP

[6]This type of econometric model is frequently utilized in analysis with international trade data to account for shocks at the country-year level and industry fixed-effects. The procedure with both county-and-município and state-year fixed effects is extremely demanding on the data.

[7]https://www1.folha.uol.com.br/mercado/2020/02/informalidade-atinge-recorde-em-19-estados-e-no-df-diz-ibge.shtml

counties also overlap in key characteristics, allowing for direct comparisons. These are the only two large countries, to the best of my knowledge, with this type of data readily available. I argue these are the best data to test these relationships since it would be difficult or impossible to validate the relationship in areas with poor-quality GDP data.

Building on an important finding in the nighttime lights economic growth literature that physical characteristics contribute significantly to economic growth, I collapse all observations to their county and município-level means a between-county or between-município procedure (Henderson et al., 2018). Rather than using the soil toposequence and other physical characteristics, I depart from the list of characteristics tested by Henderson et al. (2018) and consider different physical characteristics as well as public goods, in particular, those areas where concentrated economic activity takes place. I include important elements of infrastructure that, *a priori*, are known to be centers of economic activity. Included in the model are: ports, border crossings, airports, roads, railways and navigable waterways. If these elements contribute significantly to nighttime light, controlling for population and GDP, that demonstrates the importance of accounting for physical characteristics and infrastructure elements in estimating models that incorporate nighttime lights at a high geospatial resolution. If the dummy variables are significant contributors to light, it reinforces the importance of these attributes and characteristics as centers of concentrated economic activity. In general, the omission of public goods such as ports, border crossings, airports, roads, railways and the presence of navigable waterways from cross-sectional regressions can be problematic as this may drive both economic growth (GDP) and nighttime light. Again, the between-county, between-município approach is useful in that it allows me to estimate the marginal contribution to light of certain infrastructure elements or physical characteristics despite their time-invariant nature.

With respect to the burgeoning literature of economics papers using nighttime light, much of the applied economics literature to date has utilized the older generation of satellite images, the DMSP satellite data. Unless otherwise noted all of the following papers utilize DMSP rather than VIIRS data. Authors Pinkovskiy and Sala-i Martin (2016) use nighttime lights to evaluate the relative quality of national accounts data over household survey data. Jedwab et al. (2017) examine path dependence manifested by the establishment of colonial-era railways and the effects of colonial railways on modern day development in Kenya. The authors in Jedwab et al. (2017) use nighttime lights as their measure of contemporary economic development. Conceptually similar to Jedwab et al. (2017), another paper examines the persistent effects of Roman roads on contemporary economic development in Europe (Dalgaard et al., 2018). Keola et al. (2015) analyze growth in developing countries using nighttime lights. The authors propose that nighttime lights may not be able to distinguish economic activity in agricultural areas where light may not scale with productive activities. Michalopoulos and Papaioannou (2013) investigate pre-colonial institutions and explore how they shaped regional economic development using nighttime lights as an indicator for economic development. The authors find a strong correspondence between pre-colonial institutions and present-day economic development. Similarly Ranjan and Talathi (2021) examine the effect of colonial institutions on present-day economic development in India using nighttime lights to measure contemporary economic growth. Mirroring findings in other papers, the authors conclude that areas less impacted by colonial institutions grow more rapidly, though there appears to be convergence (Banerjee and Iyer, 2005). Cook and Shah (2020) analyze the effects of India's rural employment guarantee program using nighttime lights and find evidence for beneficial economic effects of the program.

Gennaioli et al. (2013) take a deep dive into the roots of regional development by testing for a correspondence between human capital and regional development, though they use nighttime lights as a robustness check rather than as a primary method. Jean et al. (2016) use nighttime lights and machine learning to create a model for predicting poverty at a highly disaggregated level. Michalopoulos and Papaioannou (2014) use nighttime lights to estimate the effects of ethnic divisions and institutions on economic outcomes. The authors find that institutions do not fully explain differences in within-ethnic group economic outcomes. Alesina et al. (2016) use nighttime lights to measure the effects of different geographical endowments on economic well-being. The authors identify the presence of an inverse relationship between contemporary economic development and ethnic inequality.

Baum-Snow et al. (2017) explore how railroads and highways have influenced the Chinese urban landscape. In their paper, railroads and highways are found to displace populations in China and, the authors argue, may create a negative effect by decentralizing economic activity. Henderson et al. (2018) explore whether geography influences the spatial distribution of human economic activity proxied by light. The authors find that geographic characteristics account for as much as 50% of the variation of economic activity (light). In less-developed countries the authors find that agricultural contributions explain more variation in light than do changes in international trade. Gennaioli et al. (2013) evaluate regional development and convergence using a new dataset of regional GDP and cross-validate their findings with night lights data. Henderson et al. (2017) attempt to identify the causes of urbanization in Africa utilizing nighttime lights data. The primary hypothesis of Henderson et al. (2017) is that urbanization may be shaped by climate change as a primary force.

Hodler and Raschky (2014) examine the presence of stronger contemporaneous growth in regions or states associated with the leader of a country and find a significant result. The authors conclude that during the term of a premier the region from which that premier comes enjoys higher GDP growth in relation to the rest of the country. Mellander et al. (2015) examine the relationship between economic activity, population, enterprise density and nighttime light in Sweden. Utilizing high-resolution geospatial data on enterprises and enterprise characteristics, the authors find that light growth corresponds most strongly to nighttime population density (population) rather than daytime enterprise density. A significant limitation of the analysis in Mellander et al. (2015) is that the authors use cross-sectional rather than panel data. Mellander et al.

(2015) 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, though the effect of GDP on nighttime light in the United States is estimated to be 0.636 with my preferred specification (table 2) that incorporates state-year dummies. Levin and Zhang (2017) utilize data from the newer VIIRS satellite and analyze 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.

One recent 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. There appear to be no permanent effects of flooding on city development (Kocornik-Mina et al., 2020). The authors utilize the prior generation of nighttime lights to measure economic recovery from large-scale floods in over 1,800 cities across 40 countries. The authors find that low-elevation areas are more likely to flood, and they are also fast to recover from damage. Low-lying areas are centers of concentrated economic activity, and the authors find no evidence that economic activity endogenously relocates to higher, more secure areas. This work represents one of the strongest examples of the type of analysis that can be done with nighttime lights.

Bluhm and Krause (2018) use nighttime lights images to measure primate cities in sub-Saharan Africa and the growth of primate cities.[8] The authors highlight the potential benefits of sub-national or regional measurement of economic activity using lights and offer some critiques of the shortcomings of the DMSP technology. The primary purpose of Bluhm and Krause (2018) is to document the increases in the size of primate cities and test if city lights follow a pareto distribution. Frick et al. (2019) use DMSP 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

[8] A primate city is very large primary urban agglomeration that is the social, economic and legislative center of a country

zones were links with pre-existing industrial infrastructure in the surrounding area and the presence of large markets in which to sell outputs. Bleakley and Lin (2012) use night-time lights from the years 1996-7 to test for path-dependence around natural water features in the United States such as waterfalls. The authors find that portage sites, locations 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) leverage the global nighttime lights coverage to estimate the fraction of the population below the poverty line. They find that spillovers from economic activity rarely reach to rural populations. Similarly Bruederle and Hodler (2018) use DMSP lights and find that nighttime light is a meaningful proxy for economic development at the local level in sub-Saharan Africa. Asher et al. (2021) use DMSP lights to test for correspondence between DMSP-measured nighttime light and village-level characteristics such as the population, employment, per capita consumption and electrification though they lack data on output perhaps due to the presence of a large informal sector in most villages.

In partnership with the GiveDirectly charity, a recent paper by Egger et al. (2019) utilized around 10,000 households in 653 villages to make unconditional cash transfers of a sizeable amount. This randomized controlled trial was designed to estimate the size of the famous Keynesian multiplier. Building on the back of that experiment, Huang et al. (2021) analyze the relationship between local development and nighttime lights though using VIIRS data rather than DMSP data. The authors find that NTL may be a poor proxy in some rural areas of Kenya, though nighttime lights are found to capture some fine-grained changes in rural areas. Gibson et al. (2021) outline the reasons for preferring the VIIRS series to the DMSP nighttime lights and tests for a relationship between economic output and nighttime light in Indonesia though in their context the authors use nighttime lights as a predictor rather than the dependent variable. They find a persistent relationship which is even stronger with VIIRS nighttime lights compared to DMSP. The authors demonstrate VIIRS lights better capture the rural/urban split relative to DMSP nighttime lights.

Gibson and Boe-Gibson (2021) analyzes the relationship between county-level GDP in the USA and a combination of VIIRS/DMSP data starting in 2001. The authors include an element of breaking down estimates by different population densities, though there is no panel-data within-county analysis of the strength of the population-lights relationship (elasticity) or the partial relationship between GDP and VIIRS nighttime light. Unfortunately the DMSP data are known to display "blooming" or "bleeding" effects where light seeps from one pixel into the next as measured at the sensor (Hao et al., 2015). VIIRS imagery allows resolution sufficient to address this important question at the within-county level. Similarly (Bluhm and McCord, 2022) examines the county-level elasticity of DMSP nighttime lights with respect to GDP. The authors incorporate a dimension of population into their estimates though they do not directly estimate the partial effect of population on nighttime light.

Two other recent studies demonstrate the utility of VIIRS data for economic analysis at high geospatial resolutions. Chen and Nordhaus (2015) combine nighttime lights, DMSP lights and data on output and population from Kenya. The author's data are much briefer and there are no estimates for partial correlations. Another concern with their approach is that the author's model did not employ village-level or grid-cell fixed effects that would account for local-level unobserved time-invariant heterogeneity such as the presence of unobserved public goods. A different study based on gross regional product data from Chinese counties finds a close relationship between VIIRS lights data and county-level GDP though the authors do not include population changes as part of their estimates (Li et al., 2013). In general, BEA/ACS/IBGE data are considered to be of a high quality and therefore I am able to leverage the full USA and Brazilian samples incorporating population size estimates and employ a multi-level fixed-effects estimation technique.

I view this paper as being related most closely to or a successor of Henderson et al. (2012), though the paper is most closely related to Mellander et al. (2015) and Hu and

Yao (2021). My analysis marks a substantial increase in the resolution at which GDP-lights relationship is tested and validated. Henderson et al. (2012) did use panel data techniques, though with the older generation of nighttime lights while Mellander et al. (2015) used VIIRS lights at a high resolution in Sweden, but no panel data. Hu and Yao (2021) find evidence of nonlinearities and these findings are directly incorporated and tested in my analysis. This paper thus fills a needed gap and clearly demonstrates that light is a reasonable proxy for GDP under certain conditions. It is my hope that this paper can set the stage and provide support for many more nighttime lights papers to come.

2 Methodology

The main approach of this paper is to use panel-data econometrics to accurately measure the elasticities and to decompose the links between population growth, income growth and nighttime light as measured. Using nighttime light as the dependent variable makes sense in the context because population is included as an independent variable. In this case it does not seem logical to include nighttime light as an independent variable and GDP as the dependent variable. It would not make sense to decompose GDP into its constituent components: nighttime light and population, for example. The satellite images from the VIIRS may be noisy even after processing, and it is better to use a potentially noisy variable as the dependent variable such that measurement error may have less influence on the estimation of the coefficient. Despite minor drawbacks the images are very precise in how they record the texture of activity across space as depicted in the figures 1-3 below (Chen and Nordhaus, 2011).

Given the density of counties and municípios and their explicit spatial relationship, it is critical to incorporate controls for spatially-correlated economic shocks using the procedures developed by Conley (1999), Conley and Molinari (2007) and applied by Hsiang (2010) among other applied econometrics papers. This method uses a non-parametric bootstrap estimator of the covariance to account for the underlying spatial structure of the data. The general model states simply that night-time light is a function of income, population and other factors:

$$NTL_{ct} = \beta[GDP_{ct}] + \alpha[POP_{ct}] + \gamma_c + \psi_{st} + \varepsilon_{ct} \tag{1}$$

Where c indexes the county or município, t indexes the year, γ_c are the county/município fixed effects and ψ_{st}, state-year fixed effects, which control for *time-variant*, unobserved, state-year specific economic shocks such as price shocks, political elections or other economic volatility including weather shocks in addition to the county-level controls for county-and-município-level unobserved heterogeneity. Though computationally expensive, I argue these results allow the most robust and precise estimates of the effect of GDP on lights. I use the sum of light pixels within a given county or município, NTL_{ct}, (with radiance measured in nW/cm^2/sr), as the measure of nighttime light in a given county in a given year. All continuous variables have been log transformed such that the resulting estimate is the elasticity in percentage terms.

Based on previous papers such as Hu and Yao (2021), there is reason to believe that income and population may not enter the nighttime light production function linearly. This is an important consideration for our purposes as nonlinearities may mask the effects of interest. For these reasons I also estimate an alternate specification that includes squared terms and interaction terms as independent variables. The intuition behind the squared terms is that there could be strongly diminishing effects of income and population on nighttime lights. The interaction term is included to capture the possibility that the lights-income and lights-population relationship could be amplified (or dampened) in more populated, wealthier counties and municípios. The second potential specification is therefore the following:

$$
\begin{aligned}
NTL_{ct} = \beta[GDP_{ct}] + \beta[GDP_{ct}]^2 + \alpha[POP_{ct}] + \alpha[POP_{ct}]^2 + \\
\alpha[POP_{ct}] \times [GDP_{ct}] + \gamma_c + \psi_{st} + \varepsilon_{ct}
\end{aligned} \tag{2}
$$

Included in the estimates are the log-transformed variables, the squared transformation of the control variables and the interaction of the log of GDP and population. All primary specifications include state-year fixed effects.

Between-county Estimation

There are geographic and physical characteristics of counties and municípios which we may like to analyze but the presence of infrastructure features are largely invariant within the sample period of 2012-2019. The effect of infrastructure and other time-invariant features are therefore “washed out” by the within-county or within-município fixed-effects procedure. Given the size of the sample and the survey period, using the between estimator is an appropriate approach to consider the marginal effects of geographic variables on nighttime light. In order to obtain identification of time-invariant features, all variables are collapsed to their county-level means. Identification of the effect of the infrastructure or geographic features then comes from comparing between counties which have infrastructure or features to other counties within the same state that lack infrastructure features. As the sample period is short, I argue the presence of infrastructure elements is unlikely to be endogenous to nighttime light or GDP within the sample period. Roads, airports, rail lines and ports for the most part would have already been present at the start of the sample period (2012-2019) or in many cases these features take many years to prepare and construct. The estimated equation using the between estimator is:

$$\begin{aligned}\widehat{NTL_c} = \beta[\widehat{GDP_c}] + \alpha[\widehat{POP_c}] + \phi_1[\text{PORT}_c] + \phi_2[\text{ROAD}_c] + \phi_3[\text{AIRPORT}_c]+ \\ \phi_4[\text{RAIL}_c] + \phi_5[\text{BORDER}_c] + \phi_6[\text{Water}_c] + \psi_s + \varepsilon_c\end{aligned} \tag{3}$$

where the hat refers to the county and município-level means of GDP and population over the years 2012-19.

3 Data

I contrast data and estimates from the United States and Brazil, two countries which have some similar characteristics and some differences. The United States and Brazil were the two largest countries for which annual data on both population and GDP were readily available at a high geospatial resolution in the years of operation of the VIIRS. The two countries combined make up 6.7% of the global population. Both countries are stable democracies, one very wealthy in the USA, and one country which has a substantial degree of income inequality and, presumably, wealth inequality in Brazil. The 3,095 counties of the United States provide a large landmass and total population to use for testing the nighttime light-GDP-population relationship. The United States enjoys substantial heterogeneity with respect to landmass, as well as demographic composition and population density. Both Brazil and the United States feature diverse geographic characteristics including mountains, lakes, rivers and coastlines as well as vast networks of infrastructure. The differences within the United States are evident when considering places like California, with only 58 counties per 40m citizens; Alaska which has substantial oil wealth, enormous counties and extremely tall mountains, though it 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 while Hawaii is a tropical island halfway between the US and

Japan in the middle of the Pacific ocean.

A country with 211 million people[9] living in 5,570 municípios, Brazil is also diverse in environmental and geographical characteristics. Though municípios are, on average, smaller than counties, there is significant overlap between município size and county size. There is also substantial heterogeneity in the geography of Brazilian municípios ranging from the unique coastal city of Rio de Janeiro to Manaus in the middle of the Amazon rainforest. Brazil has dense and poor areas to a much larger extent than the USA. Since the two countries combined include many heterogenous county and município types I analyze the USA and Brazil separately, but present the results of comparable estimates side-by-side in the text.

Table 1 details years of data availability. The VIIRS nighttime lights series starts only in 2012 while GDP data at the county level are available from 2001-2020 for the US and for a similar period for Brazil. County-level population estimates for the U.S. start in 2009 and are available until 2020. This analysis is therefore limited by the lack of current GDP data from Brazil as we have no GDP estimates at the município level past 2019 for Brazil.[10]

		Source	Years Available
GDP	USA	BLS	2001-2020
	Brazil	IBGE	2002-2019
Population	USA	ACS/census	2009-2020
	Brazil	IBGE	1975-2020
Lights	Both	NoAA/NASA	2012-present

Table 1: Data Availability

3.1 BLS/IBGE GDP Data

Over the past years the Bureau of Economic Analysis (BEA) at the U.S. Bureau of Labor Statistics (BLS) has released local-area calculations for gross domestic product.

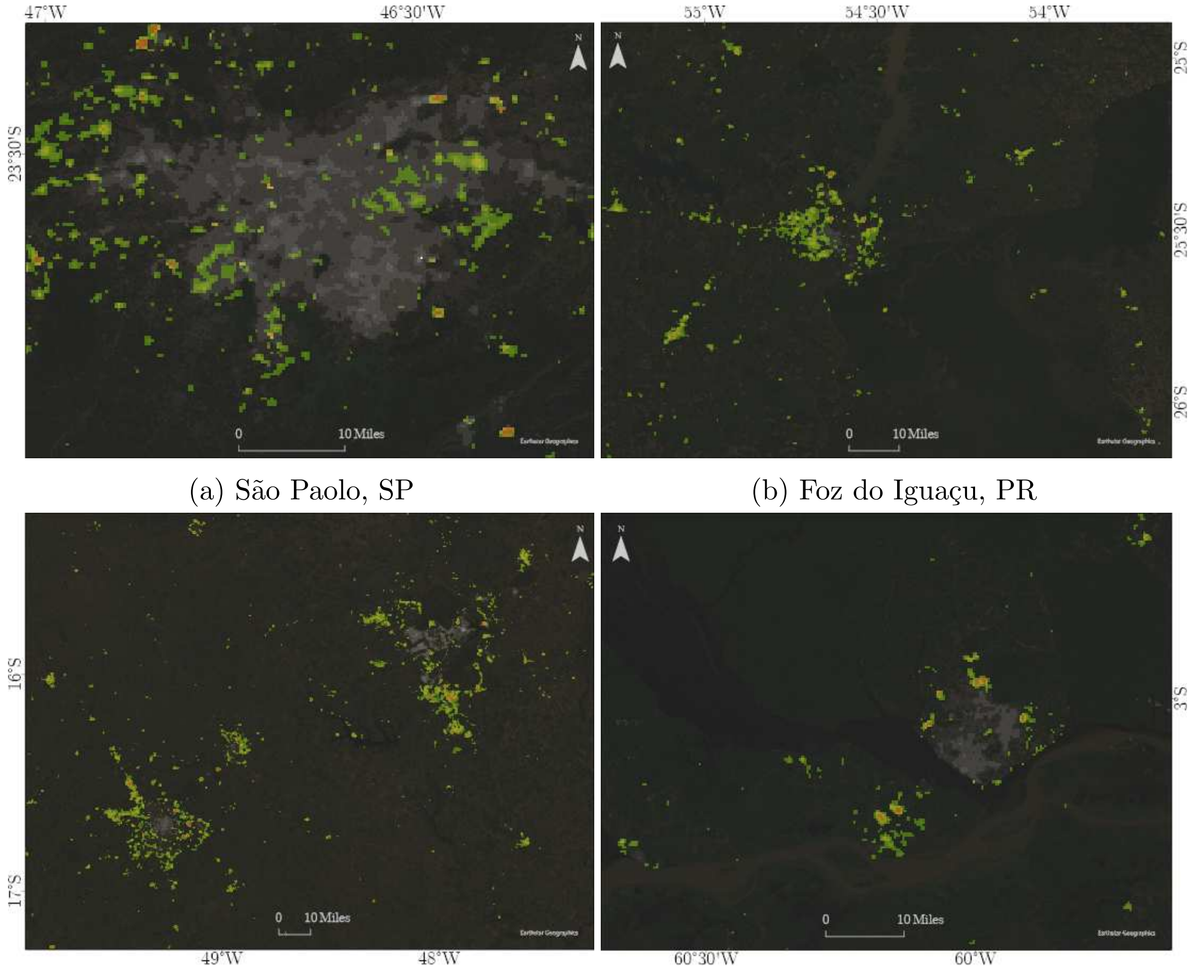
[9]Source IBGE Census Data: https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?=&t=resultados

[10]Although 2020 USA data are available at the time of this writing, for consistency I have elected to use only data from 2019 and earlier. At the time of this writing, 2020 GDP estimates are not available for Brazilian Municípios.

In the BEA/BLS GDP statistics county-level GDP is calculated using the income approach. Based on the availability of data the 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).[11] There is substantial between-county variation in the GDP data as some counties produce output worth millions of dollars while others produce well under 100k per annum. The Brazilian GDP data comes from the Instituto Brasileiro de Geografía e Statística (IBGE). The data are compiled from governmental and other administrative data sources, similar to the U.S.A. GDP estimates.[12]

[11] Principal sources of the county-level GDP data are the Department of Labor's Quarterly Census of Earnings and Wages, air-carrier traffic statistics, Department of Transportation surface transport 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).

[12] The full details of all sources and methods for the production of the Brazilian GDP estimates can be found on the IBGE website: https://biblioteca.ibge.gov.br/visualizacao/livros/liv97483.pdf accessed Feb-17-22



(a) São Paolo, SP

(b) Foz do Iguaçu, PR

(c) Brasilia, DF

(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 2012-2019 - Green = small change, Red = large change **(TO BE PRINTED IN COLOR)**

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

Population estimates come from ACS 5-year estimates of the county-level population. These are calculated using data sampled from counties on a rolling basis over the course of 5 years. ACS data are the main survey data for inter-censal periods. Like the GDP estimates the Brazilian population estimates also come from the IBGE. The estimates are based on the Brazilian population census which took place in 2000 and 2010 and adjusted for changes in between.

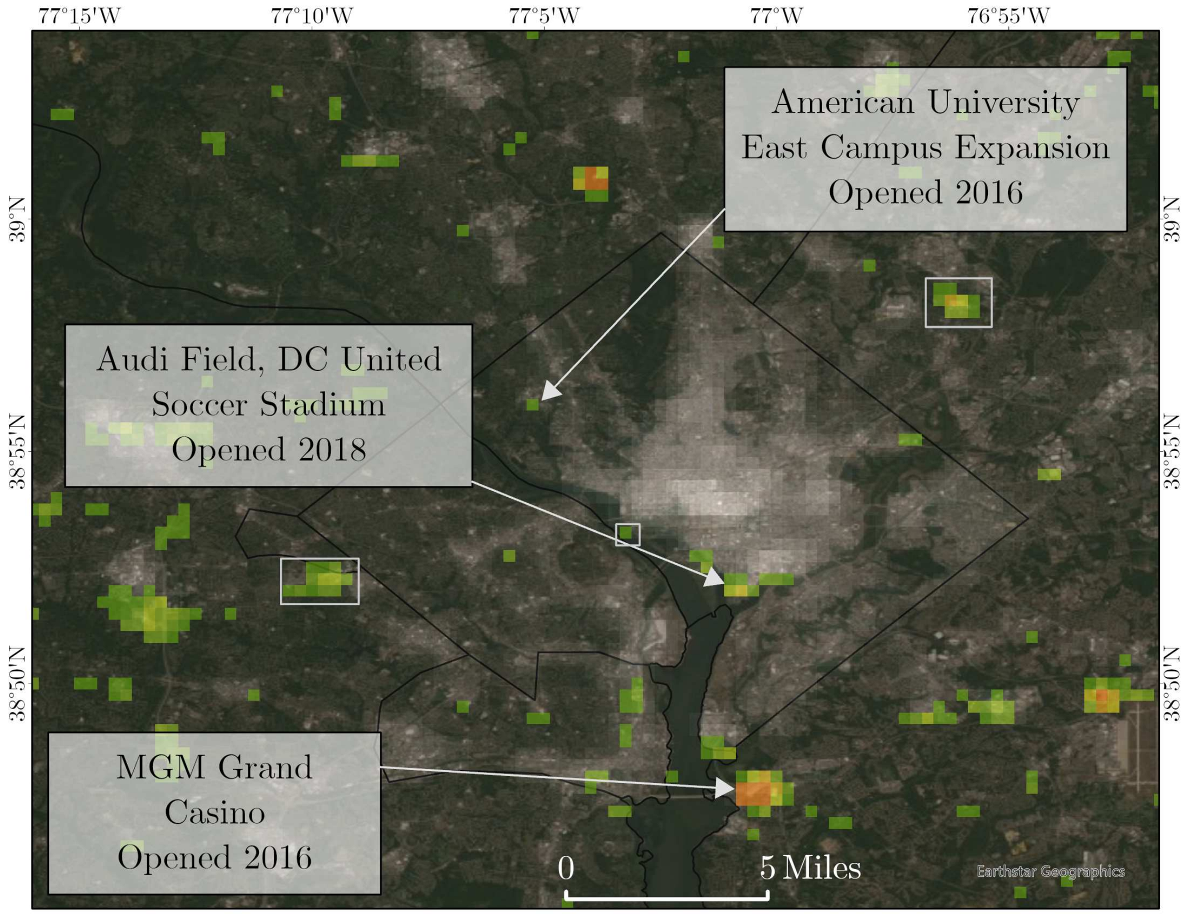


Figure 2: Night-time growth in Washington, DC 2012-19. Green = small change, Red = large change. Contemporaneous daytime imagery of economic development in the boxed areas is shown in the appendix. Layers: Basemap: Open Street Map, CC License; Night-time Lights Annual Image (2019) **(TO BE PRINTED IN COLOR)**

3.3 VIIRS Night-time Lights Data

The Visible Infrared Imaging Radiometer Suite (VIIRS) is designed to capture human-made light and overcomes many limitations of the previous Defense Meteorological Satellite Program (DMSP) satellite images. The images have been collected since 2011 for a joint partnership between the National Aeronautics and Space Administration (NASA) and the National Oceanographic and Atmospheric Administration (NOAA) and are hosted by the Earth Observation Group based the Colorado School of Mines. The latest high-resolution images of the earth at night are captured on-board the

Suomi-NPP satellite every night using the VIIRS.[13] VIIRS nighttime lights images have resolution 45 times higher than the previous generation of nighttime lights images which had ground footprint[14] of 5km by 5km ($25km^2$) while VIIRS ground footprint is a mere 742m by 742m or $0.55km^2$ (Elvidge et al., 2013).[15] The VIIRS incorporates an automatic gain sensor which adjusts allowing greater sensitivity and reducing the need for performing calibration procedures with the images. This also means the sensor can better capture much lower and higher levels of light than the previous generation (Elvidge et al., 2017).[16] The automatic gain sensor also attenuates limitations around night-time lights data coming from heavily saturated urban areas. The new VIIRS images are available on a daily frequency or in monthly composite forms as well as some yearly composite images (Carlowicz, 2012). The enhanced resolution is of interest to researchers attempting to pinpoint precise centers of economic activity. An additional advantage is that the VIIRS ground footprint does not expand further from the nadir unlike the previous generation of satellites. This leads to increased precision away from the nadir (Chen and Nordhaus, 2015).[17] For the data estimated here I utilize the V1 monthly VIIRS nighttime images compiled into annual composite images using a

[13]https://www.nasa.gov/mission_pages/NPP/main/index.html

[14]ground footprint of the satellite is the resolution of the output product. The pixel size of the satellite is smaller though pixels are blended together to save memory. This also introduces geolocation errors as outlined in (Abrahams et al., 2018) and (Tuttle et al., 2013) though the presence of geolocation errors is greatly mitigated with the VIIRS sensor suite. VIIRS images also suffer less "blooming" than DMSP meaning blurring of the light image.

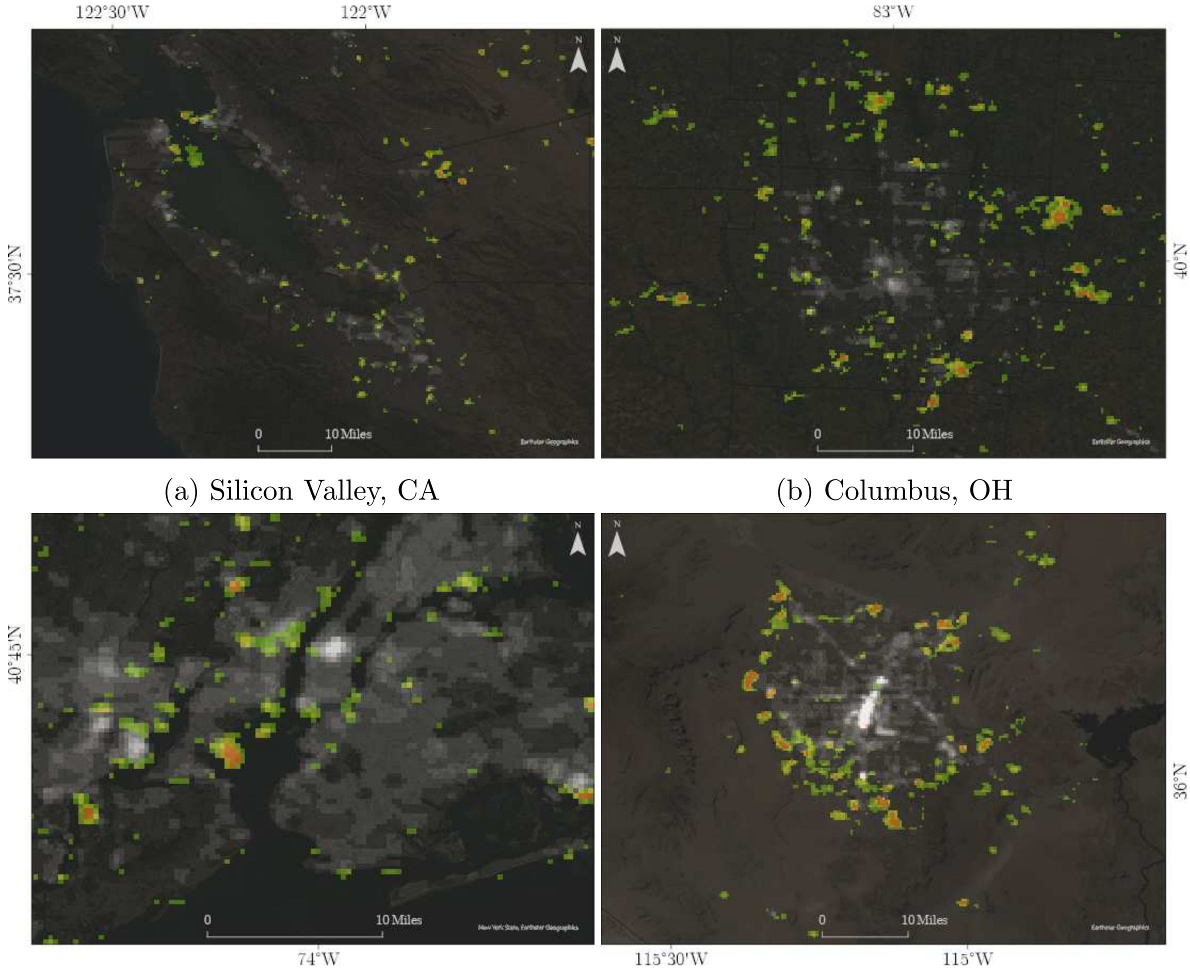
[15]With regard to the older generation of satellite images there were several known issues including a wide margin of precision. DMSP satellites identified pixels with a margin of error of 2.9 km, as tested in Tuttle et al. (2013), and this margin of error appears to be much smaller with the VIIRS data (see figure 2). Due to the technology on the DMSP satellites and the fact that the earth is viewed at an angle, the DMSP images ground footprint increases as one gets further from the nadir of the satellite. This improvement in precision menas the VIIRS images do not face the same limitations as DMSP. It is possible to leverage worldwide VIIRS data for some analysis though it would seem prudent to be cautions when combining data across countries.

[16]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. Light is then measured in radiance with the unit being $nW/cm^2/sr$. A detailed accounting of the initial processing of the data can be found in Elvidge et al. (2017).

[17]The nadir of the satellite is the point at which the satellite is furthest from the earth.

weighted sum.[18] I include estimates using nighttime lights data that has undergone further processing, masking as well as gas flare removal, to remove residual light features unrelated to human economic activity. Those estimates, where applicable, are included in the appendix. In general, I find that there are no substantial advantages to using masked versus unmasked data or gas-flare-removed data.

[18]Details of how this procedure was accomplished are included in the online appendix section E



(a) Silicon Valley, CA

(b) Columbus, OH

(c) New York City, NY

(d) Las Vegas, NV

Figure 3: Night-time Lights of Four Major American 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 **(TO BE PRINTED IN COLOR)**

Some examples of night-time lights images of major Brazilian cities and U.S. cities are shown in figures 1-3. Long-run changes in night-time light are shown in green-red colors to demonstrate intensity. Figure 1, panel (a) is São Paulo, SP which is by far the most populated Brazilian state at 48.6m persons. Around São Paulo there appears to be substantial development and sprawl especially along the coastline and the highway corridor. In panel (b) of figure 1 the city of Foz do Iguaçu, PR, Brazil is visible where the Itaipu hydroelectric dam straddles the border with Paraguay to the East and Argentina to the South. Differences in economic development are apparent on the Paraguayan side relative to the Brazilian side demonstrating the sensitivity and high-resolution of the VIIRS sensor. 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 bottom left corner of the figure, panel (c) shows Brasilia, DF with economic growth visible down to Goîana in the bottom left corner with the city of Anápolis in between. This area has experienced a relatively rapid period of development compared to other parts of Brazil. In panel (d) we have Manaus, a city in the middle of the Amazon rainforest. In Manaus the increases in the intensive margin, light intensity, are clearly much more intense than changes in the extensive margins that would correspond to outward expansion of nighttime light.

Figure 2 demonstrates the resolution of nighttime lights and the fine-grain detail of economic development that can be clearly seen. In Washington, D.C., despite high density of lights, changes in light intensity can still be distinguished at a high resolution. The dark red spot just south of Washington, D.C. is the MGM grand casino, nearly always lit, and an area of major economic development for the D.C. metropolitan area over the last few years. Another major development in D.C. over the same period was the construction of a new soccer-specific stadium in the Buzzard Point neighborhood. Stadium plans had been in development since 2014, though the team had been searching for a stadium site for years prior to the Buzzard Point location. The stadium is glowing yellow dot where the Potomac river meets the Anacostia river at the southern tip of diamond-shaped D.C.

In figure 3, panel (a), Silicon Valley, 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. Panel (b) shows Columbus, Ohio, one of the fastest growing areas in the United States in recent years. In Columbus, economic development is quite dispersed in comparison with Las Vegas, for example. In New York City, the most populated city in the country, Times Square is clearly visible in the lower central area of Manhattan. A majority of economic activity is taking place in the docks/port at Newark, which is part of the metropolitan area though those two areas are separate counties. Last, Las Vegas, Nevada in panel (d) makes for a striking example because of its intensity relative to the darkness of the nearby un-populated desert.

3.4 Infrastructure Data

USA infrastructure data including the location of ports, rail, navigable waterways and the location of border crossing points have been collected from the U.S. federal government's Homeland Infrastructure Foundation Level Database (HIFLD). Airport locations were taken from open data sources.[19] Data on primary roads, which includes interstates and principal highways, were collected from the US Census Department.[20] All Brazilian infrastructure data come from the Brazilian Infraestrutura Nacional de Dados Espaciais (INDE) [21] geospatial database.

4 Results

4.1 Linear and Non-linear Estimates

Table 2 contains the results of the primary model using the average radiance nighttime light data.[22] Again, these models have been fit using the Conley spatially-adjusted standard errors. The threshold distance for the spatial influence of economic shocks is set to 5,500km or 3,417 miles, roughly the width of the continental United States. The parameters are set to allow for infinite location-specific serial correlation. Table 2 column 1 are the estimates for the effect of population and GDP on nighttime light

[19]https://ourairports.com/
[20]https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2020&layergroup=Roads
[21]https://inde.gov.br/
[22]The linear model was estimated with masked, unmasked, and gas-flare-removed versions of the nighttime light products. There appears to be little difference between using the two in the case of these estimates. Results are included in the appendix table 2.

	USA		Brazil	
	(1)	(2)	(3)	(4)
	Sum of Avg. Yearly Radiance	Sum of Avg. Yearly Radiance	Sum of Avg. Yearly Radiance	Sum of Avg. Yearly Radiance
GDP	0.636***	2.096***	0.118***	1.212***
	(0.0872)	(0.194)	(0.0331)	(0.119)
Population	-0.294***	-0.934***	0.256***	0.573***
	(0.0671)	(0.253)	(0.0401)	(0.128)
GDP^2		-0.0477***		0.00436
		(0.0170)		(0.00664)
Pop^2		0.100***		0.0711***
		(0.0188)		(0.0132)
GDP×Pop		-0.0709**		-0.127***
		(0.0331)		(0.0166)
Observations	24,670	24,670	44,547	44,547
State-Year FE	Yes	Yes	Yes	Yes
Spatial Kernel	5500km	5500km	5500km	5500km

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

Conley spatially corrected standard errors in parenthesis.

All columns contain County/Município and year fixed effects.

Table 2: Nighttime Lights Regressions with State-Year Dummies

for the USA. The partial effect of GDP on nighttime light is estimated to be strong, positive and statistically significant while the partial effect of population is estimated to be negative. The estimates for the GDP effect, β, are consistent with the intuition that a greater amount of output corresponds to a greater amount of light, the effect of population goes against the intuition that each individual consumes a certain amount of light, and thus it seems logical that light should be *increasing* in the size of the population holding GDP constant. Column 2 contains the estimates of the nonlinear transformed model. It is important to note that these are average marginal effects. The marginal effects plots can be found in figure 4. Again we have the positive (expected) sign on the GDP variable, though the relationship between population and nighttime light is once again estimated to be negative. The squared transformation variables are interesting because the sign on the effect of GDP^2 is negative, meaning the (log of) the sum of nighttime light is increasing in GDP/output at a decreasing rate. Looking at the marginal effects in panel (a) the effect of GDP on nighttime light is shown to become negative around 14 log-points of population. With respect to the population effects, we see a very different story where the effect of population starts out negative at the bottom of the population distribution of the USA and climbs to turn positive around the middle of the distribution at 10 log points.

The Brazil estimates strike a meaningful contrast with the USA estimates. For estimates with the Brazilian data, in columns 3 and 4, the effect of both GDP and population are strictly positive; the effect size is increasing in both population and GDP. Most importantly, the effect of population on nighttime light is stronger in Brazil relative to the effect of GDP. This has substantial implications as, in Brazil, nighttime lights data appear to capture changes in the population better than changes in GDP.

These changes in the relative effect size reveal the complexities in the GDP, population, nighttime lights nexus and, I argue, emphasize that it can be problematic to combine nighttime lights across countries when proxying for GDP unless researchers also are prepared to be proxying for population growth in the case that no population estimates are available. In order to better understand the relative effect size between GDP and population, β & α, I decompose these estimates for different slices of the distribution in the following section.

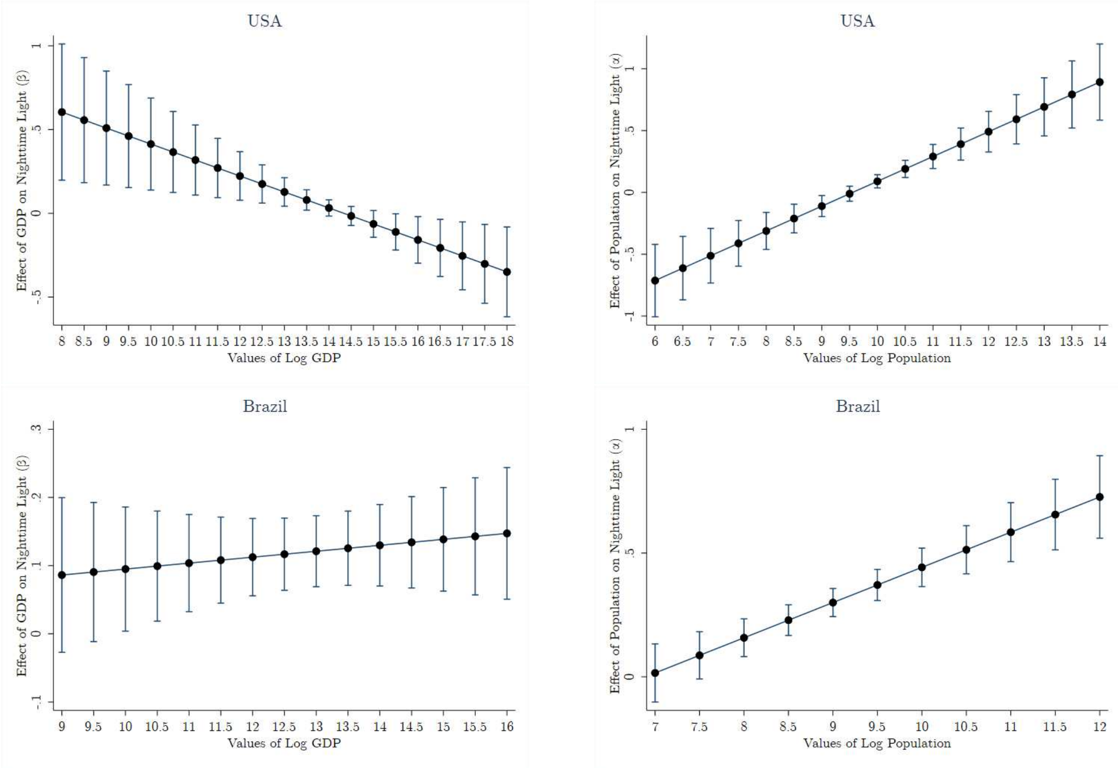


Figure 4: Marginal Effects

4.2 Regressions by Quantiles

The following analysis of the effect of population and GDP, α and β, on nighttime light divides the sample into quantiles of GDP, population and area. In each case the thresholds are standardized and estimates can therefore be compared from the lowest-income Brazilian municípios with the poorest USA counties. Table 3 compares the quantiles of counties to municípios and reveals differences in the distribution of counties and municípios. U.S. counties tend to be larger, wealthier and less populated while Brazilian municípios tend to be small and highly populated. In all estimates the results are split into the USA sample and the Brazilian sample for analysis.

4.2.1 Quantiles of GDP

Looking at the estimates by quantile of GDP in the top half of table 4 we can see the estimates for the GDP and population coefficients with columns corresponding to the individual quantiles. In the USA, across all quantiles the GDP effect dominates

	Areal Quantile	Number of Counties	Avg. Size (in Sq. km)	Avg. GDP (in USD)	Avg. Population
U.S.A.	1	18	149	67,000,000	555269
	2	112	455	5,469,884	105285
	3	717	951	3,756,954	75765
	4	1209	1616	4,513,667	91387
	5	1031	6373	6,883,178	128229
	Areal Quantile	**Number of Municípios**	**Avg. Size (in Sq. km)**	**Avg. GDP (in USD)**	**Avg. Population**
Brazil	1	1714	145	633,753	22813
	2	1619	377	930,825	33010
	3	1014	824	1,250,366	45341
	4	522	1581	1,953,698	59880
	5	700	8542	1,329,080	49148

Table 3: Characteristics OF Counties vs. Municípios by Quantiles of Size

the population effect. Interestingly at the lowest quantiles of GDP we see a positive effect of population on nighttime light for the USA. Looking at Brazil there is a very different picture. In the first 4 quantiles, the effect of GDP is estimated to be smaller than that of population for Brazil, however in the top quantile of GDP we see the relative magnitude of the effects flip, and the effect of GDP on nighttime light is larger than the effect of population. This reveals an issue where, for some parts of the distribution, the effect size is different, and in this case those changes appear, at least in some way, to be related to the level of GDP of the município. When the relative effect of population on nighttime lights vs GDP on nighttime lights changes, this could be highly problematic for researchers wishing to use nighttime lights data across different sections of the distribution of GDP.

4.2.2 Quantiles of Population

The bottom half of table 4 shows the results of estimates by quantiles of population. For the USA estimates, the size of the GDP effect dominates the size of the population for the lowest quantiles of population, though for quantile 4 and 5, the effect of population on nighttime light, α appears to be larger than the effect of GDP on nighttime light. In Brazil we see a similar pattern. Though even in the lowest quantiles the effect of population on nighttime light is estimated to be smaller than the effect of increases in GDP, moving up the distribution of population the effect of population on nighttime light is estimated to be stronger and stronger. Put simply, the more people (population), the more people influence nighttime lights. For the more populated municípios the effect of increases in GDP on nighttime light is estimated to be negative while the sign on population grows stronger.

Quantiles of:			Quantile (1)	Quantile (2)	Quantile (3)	Quantile (4)	Quantile (5)
GDP	USA	GDP	0.782	1.232	1.16	0.925	0.836
		Pop	0.134	-0.613	-0.573	-0.319	-0.292
	BRA	GDP	0.269	0.273	0.304	0.336	0.446
		Pop	0.519	0.504	0.454	0.374	0.131
Population	USA	GDP	0.617	0.599	0.442	0.317	0.275
		Pop	0.298	0.207	0.358	0.476	0.439
	BRA	GDP	0.288	0.229	0.0893	-0.217	-0.338
		Pop	0.497	0.558	0.726	1.09	1.158

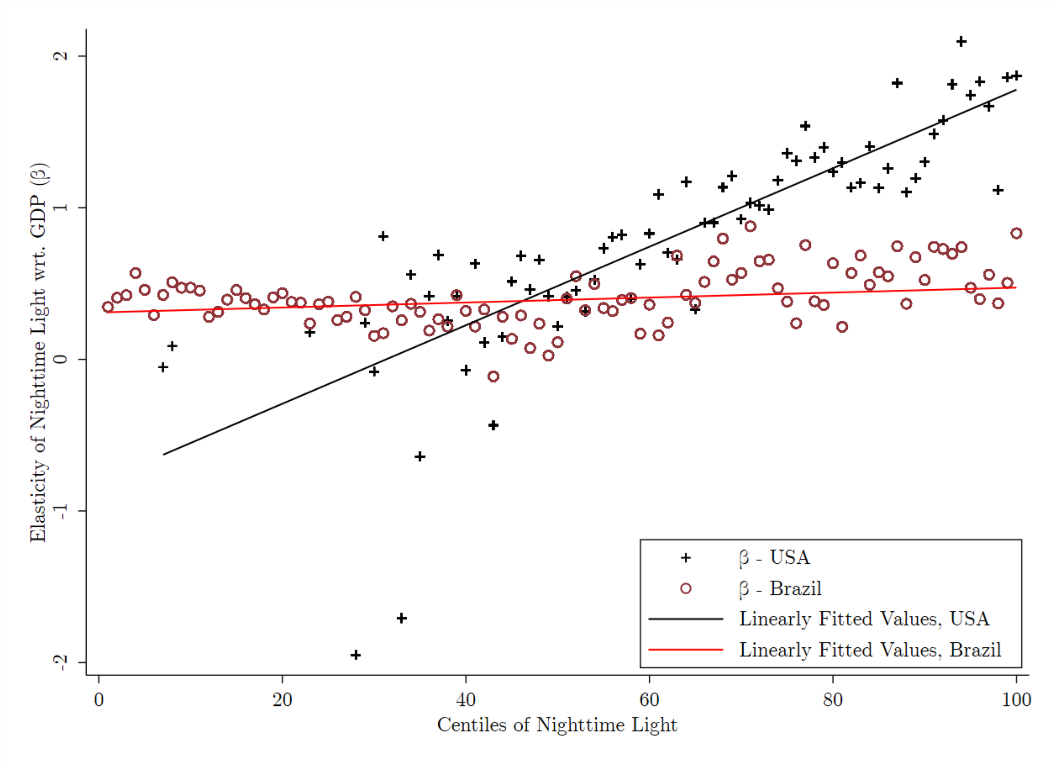
Complete regression tables with standard error estimates included in the appendix tables 8-10.
Conley spatially corrected standard errors used in all estimates.
All estimates contain County/Município and year fixed effects.

Table 4: Estimated Coefficients by Quantiles of GDP and Population

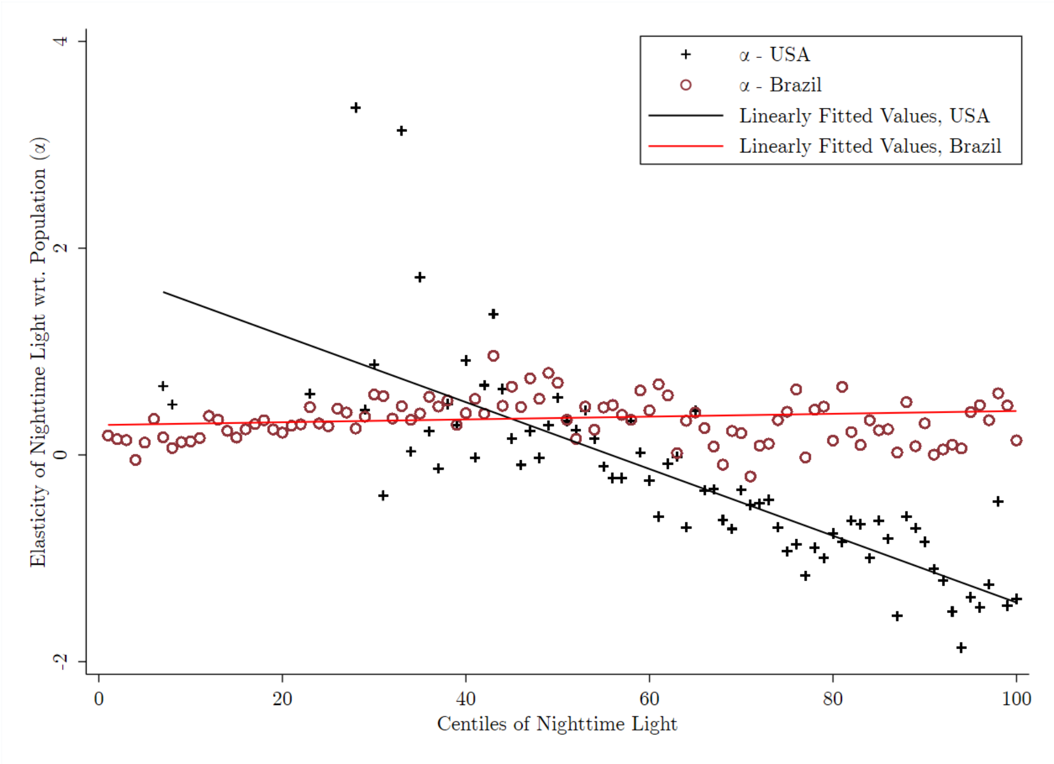
4.3 Regressions by Centile

Figure 5 examines the estimates of the effects population and GDP, α and β, on the sum of light. The distribution has been divided into hundredths (percentiles or centiles) and the model estimated separately for each centile and each country. The resulting coefficients are plotted against the centiles of light on the x-axis. In panel (a) the effect of GDP on nighttime light, β, is plotted against the centiles of light. For the USA the trend is defined and upward. In the lowest centiles the effect is difficult to distinguish from 0. Moving up the distribution to the more lit counties/municípios, the relationship or the influence of GDP on nighttime light is estimated to grow stronger. This may, unfortunately, indicate some unaccounted for variables or reverse causality. The estimates are clear, however, that the relationship is positive above a certain threshhold. With respect to Brazil, in panel (a), we can see that the effect starts out positive for the lowest quantiles and then increases slightly to the higher end of the light distribution, quite different from the USA. The effects overlap for the two countries between about the 40th centile and the 60th, so mostly in the middle of the distribution of nighttime light as measured by VIIRS.

The second panel in figure 5, panel (b), shows the effect of population on nighttime



(a) Effect of GDP on Nighttime Light



(b) Effect of Population on Nighttime Light

Figure 5

light, α, by country and by centile of nighttime light. For the USA the effect of population on nighttime light is positive until around the median of the distribution when the effect of population on nighttime light becomes negative and is estimated to be negative for the remainder of the distribution, the brightest counties. For Brazil, similar to the above with GDP the effect is estimated to be small, but positive, and increasing with the amount of light. There may be a small peak in terms of effect size around the median for Brazil, but the relationship between population size and the total nighttime light appears more or less stable across the entire distribution of nighttime light.

4.4 Economic Geography Regressions

Utilizing the capacities afforded by this data, I am able to extract estimates of the effect of infrastructure and physical characteristics on nighttime light. This is helpful as it should reveal the marginal contribution to light of particular infrastructure elements and physical characteristics. If border crossings, for example, or airports greatly increase light the presence of these elements they must not be ignored in estimation procedures and in general analyses. Estimating these effects will also give a general sense of how these resources or public goods contribute to economic development.

The economic geography variables which are included are whether the county/município has any of the following geographic or physical characteristics: the presence of a major road, the presence of a border crossing point, the presence of an airport, the presence of railway infrastructure and the presence of navigable waterways. The values of nighttime light, GDP, and population are collapsed to their county-level means for the years 2012-2019. Then the indicator variables for geographic characteristics are tested with the implied counterfactual being other counties within the same state that lack the infrastructure features. The idea behind these regressions is to capture the marginal contribution to light of each of these infrastructure elements holding income and population constant. All specifications are carried out using the same non-parametric procedure to account for spatial correlation in the error term as the previous estimates.

The results of the economic geography regressions can be found in table 5. The first columns, 1 and 2, represent regressions using USA data while the latter two columns, 3 and 4, correspond to the Brazilian data. Looking first at the parameter estimates for the effect of GDP and population on nighttime light for Brazil and for the USA, the signs of all the estimates are nearly identical to the estimates with state×year dummy variables to control for time-variant state-specific heterogeneity. One exception is the GDP×Pop interaction term for the USA, which is estimated to be negative in the within-county regressions while it is estimated as positive in the between-county regressions. This could indicate imprecise estimates, though it could also be due to the presence of state-year time-variant shocks that are controlled in the within-county within-município estimates in table 2. The size of the coefficients is of a similar magnitude to the estimates using the within-county and within-município estimates.

The primary coefficients of interest in these regressions are the ϕ_1 to ϕ_6 coefficients in expression (3). In both the United States and Brazil, the marginal effect of a border crossing on nighttime light is estimated to be strong and positive. Border crossings are obviously time-invariant and therefore their marginal contribution to light is difficult to measure outside of this type of procedure. Similar to the findings in Bleakley and Lin (2012), I believe that, like portage sites, border crossings and other geographic and physical and infrastructure characteristics are places where concentrated economic activity has taken place potentially for decades especially small border towns which host large volumes of trade flows. It also seems reasonable that these effects would be well-estimated and not endogenous. With border crossings, at least within the sample period, the presence of border crossings is not linked with changes in county or município-level GDP.

Looking at the estimates for the effects of an airport on nighttime light, the effect of having an airport is estimated to decrease light in the linear model for the USA and to increase light once the non-linear controls have been added. This type of sign-flipping of the estimates could indicate issues with heterogenous effects or the fact that airports may be endogenously added in counties or municípios with a high economic potential. In the non-linear models an airport increases light in the USA and in Brazil while in the linear models the effects of an airport on nighttime light are estimated to be negative for both countries. Having a railway or rail infrastructure corresponds to municípios and counties with less light, possibly because trains pass through un-populated areas on their way to populated ones. For the effect of rail infrastructure on nighttime light, the sign of the effect is negative across all the estimated models though the effect size is estimated to be much larger for Brazil than for the United States. The presence of a road appears to correspond to lower levels of light in the United States, while in Brazil the presence of a road indicates the presence of greater levels of light. This could indicate greater clustering around roads in terms of economic development in Brazil. This is consistent with the images included in figure 1 where development is occurring visibly along roads. In the United States there seems to be more development in pockets rather than stretched out along a road, at least from a brief visual assessment. The presence of a port is estimated to decrease nighttime light, though in both the USA estimates and the Brazil estimates we see the sign change between the linear model and the model with the nonlinear transformations. It seems intuitive that the presence of a port would increase light due to the need for processing incoming and outgoing shipments in the morning and nighttime hours. This is indeed estimated to be the case in the nonlinear models, with the effect size similar for both the Brazilian and USA nonlinear model while both estimates are unfortunately not statistically significant at standard levels. Last, the presence of a navigable waterway is associated with lower levels of nighttime light in the USA, while in Brazil a navigable water is associated with higher levels of nighttime light. In general it could be that there are many navigable waterways in the USA that are not significantly utilized for economic activity while in Brazil waterways represent a much more important route for economic activity. The fact that navigable waterways do not correspond to higher levels of nighttime light is perhaps not surprising as the presence of ports has already been included meaning that these would be areas with a navigable waterway *but no port*.

The results of this exercise point to the fact that infrastructure and physical characteristics contribute significantly to the light produced by economic and human activity. The most significant characteristic attracting additional light appears to be the presence of a border crossing, which appears to more than double the amount of light in one county or município relative to a county or município in the same state without a border crossing.

5 Test for Parameter Stability

As a test for parameter stability, although as we have seen there are some inconsistent results for different models and parts of the distribution, I drop sequentially one year's worth of data from the sample and repeat the same regressions. The results for these tests are shown in the online appendix section C. The test reveals very little change in the value of the estimated parameters for both the model with linear controls and the model with nonlinear controls. For the USA sample, the effect of GDP, β, ranges between .630 and .659 a difference of only 4%. For the effect of population in the USA, which is estimated to be negative, the effect is estimated to be between -.288 and -0.323, which appears fairly tightly estimated.

For the Brazilian sample, the estimated effect size of the effect of GDP on nighttime light is estimated to be between .0956 and .127, which is slightly larger in terms of difference as the effect size for Brazil is much larger. Between the largest and smallest estimates for the effect of GDP, the difference is 33%. For the estimates for α they fall between .245 and .285 which is a difference of 16%. All-in-all the estimates do not appear to change significantly, become insignificant, nor change sign in the case of all the parameter stability regressions.

6 Conclusion

Using precise, nationwide panel data from the USA and Brazil and 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 second administrative level. I leverage a special technique for estimating models where shocks may be spatially correlated and find that the relationship between nighttime lights, GDP and population changes is strong though the relationship between GDP and light is estimated to be much stronger than that of population and nighttime light in the case of the USA. In the case of Brazil, nighttime light appears

	USA		Brazil	
	(1)	(2)	(3)	(4)
	Unmasked NTL Average Radiance	Unmasked NTL Average Radiance	Unmasked NTL Average Radiance	Unmasked NTL Average Radiance
GDP	1.194***	3.481***	0.246**	0.221
	(0.0939)	(0.506)	(0.118)	(0.567)
Pop	-0.657***	-2.984***	0.419***	0.806
	(0.125)	(0.701)	(0.145)	(0.724)
GDP^2		-0.194***		0.0839
		(0.0406)		(0.0545)
Pop^2		-0.00584		0.116
		(0.0142)		(0.0766)
GDP×Pop		0.208***		-0.228*
		(0.0607)		(0.117)
Has Border	1.287***	0.973***	1.226***	1.409***
	(0.187)	(0.147)	(0.372)	(0.365)
Has Airport	-0.284**	0.525***	-0.0820	0.918***
	(0.120)	(0.133)	(0.217)	(0.166)
Has Railway	-0.00183	-0.0824	-0.619***	-0.338**
	(0.123)	(0.103)	(0.138)	(0.144)
Has Road	-0.585***	-0.0405	0.986***	0.121
	(0.0767)	(0.0501)	(0.370)	(0.174)
Has Port	-0.578***	0.142**	-0.332	0.128
	(0.103)	(0.0622)	(0.377)	(0.330)
Has Waterway	-0.310**	-0.164*	1.761***	1.390***
	(0.129)	(0.0918)	(0.452)	(0.334)
Observations	3,089	3,089	5,569	5,569
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Spatial Kernel	5500km	5500km	5500km	5500km

*** p<0.01, ** p<0.05, * p<0.1
Conley spatially corrected standard errors in parenthesis.

Table 5: Economic Geography Regressions

to capture changes in the population more strongly than it does changes in economic output. I believe these results provide strong evidence that night-time light changes correspond to changes in population and income at a high geospatial resolution. The relationship between nighttime light, GDP and population is strongly indicated to be different for the U.S.A. and Brazil.

The estimates appear robust after incorporating higher-order terms and interaction terms to account for the potential presence of nonlinearities in the lights-income-population nexus. Regressions divided by quantiles of the independent variables reveal distinct changes at different points in the distribution. For example, in Brazil population changes more strongly influence nighttime light than GDP changes, however the relative magnitude of those effects switches for municípios in the highest quantile of GDP. This result is important for analysts who seek to use nighttime light to test for the effects of policies on economic output or to proxy for output at a high geospatial resolution. Future researchers should pay particular attention to incorporating nonlinear terms where relevant and avoid combining nighttime lights from multiple countries particularly in cross-sectional analysis.

A between-county estimator indicates the presence of a border crossing unambiguously and substantially increases light. Other physical characteristics and infrastructure elements appear to be inconsistently estimated, perhaps due to endogenous placement of airports and roads. Areas with railways are estimated to have less light compared to similar counties and municípios without a railway. These findings are useful to future researchers looking to use VIIRS imagery for high-resolution or high-frequency economic analysis with nighttime lights.

The potential of this great tool is just beginning to be understood.

	USA		Brazil	
	Obs.	mean	Obs.	mean
Masked, Sum of Avg., Yearly Radiance	24,670	21,732.02	44,547	6,713.55
Unmasked, Sum of Avg., Yearly Radiance	24,670	22,407.38	44,547	7,519.76
Gas Flares Removed, Sum of Avg. Radiance	24,670	22,307.53	44,547	7,512.86
GDP	24,670	5,514,556	44,546	1,103,727
Population	24,670	103,063	44,547	36,666
Has Border Crossing	24,670	0.02	44,547	0.00
Has Airport	24,670	0.32	44,547	0.02
Has Railway	24,670	0.88	44,547	0.22
Has Road	24,670	0.45	44,547	0.97
Has Port	24,670	0.03	44,547	0.01
Has Navigable Waterway	24,670	0.30	44,547	0.03

†nW/cm2/sr

Table 6: Descriptive Statistics for All Regression Variables

Appendix

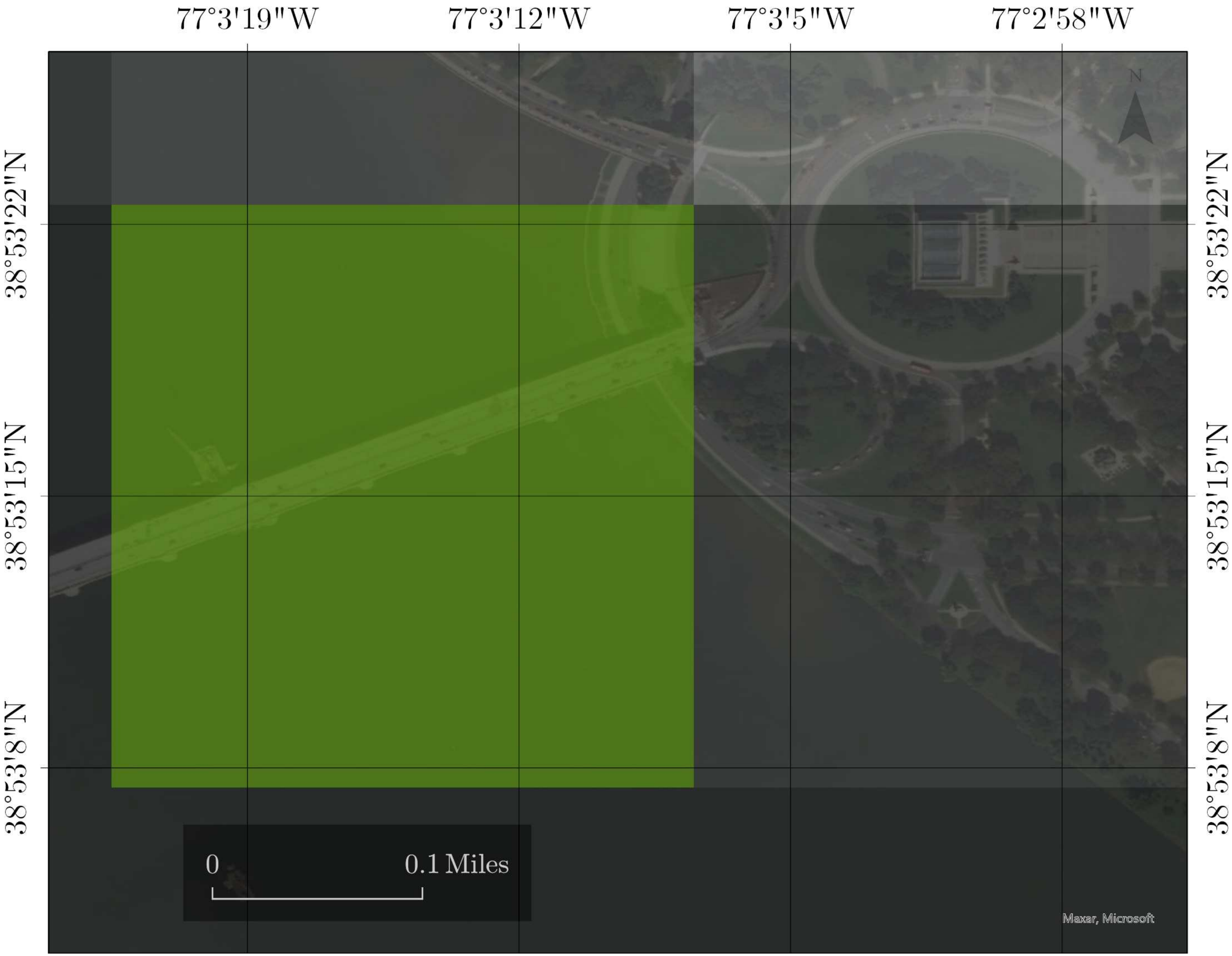


Figure 6: Night-time growth Memorial Bridge. Layers: Basemap: Open Street Map, CC License; Night-time Lights Annual Image (2019); Green = small change, Red = large change **(TO BE PRINTED IN COLOR)**

	(1) Unmasked NTL Average Radiance	(2) Unmasked NTL Average Radiance	(3) Unmasked NTL Average Radiance	(4) Unmasked NTL Average Radiance	(5) Unmasked NTL Average Radiance
GDP	0.782***	1.232***	1.160***	0.925***	0.836***
	(0.202)	(0.0393)	(0.0456)	(0.0415)	(0.0492)
Pop	0.134	-0.613***	-0.573***	-0.319***	-0.292***
	(0.292)	(0.0579)	(0.0598)	(0.0538)	(0.0638)
Observations	626	2,155	4,753	7,579	9,557
Spatial Kernel Distance	5500km	5500km	5500km	5500km	5500km

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

Conley spatially corrected standard errors in parenthesis.

All columns contain County/Município and year fixed effects

Table 7: Estimates by Quantiles of GDP - USA

	(1) Unmasked NTL Average Radiance	(2) Unmasked NTL Average Radiance	(3) Unmasked NTL Average Radiance	(4) Unmasked NTL Average Radiance	(5) Unmasked NTL Average Radiance
GDP	0.617***	0.599***	0.442***	0.317***	0.275***
	(0.0410)	(0.0601)	(0.0352)	(0.0320)	(0.0577)
Pop	0.298***	0.207***	0.358***	0.476***	0.439***
	(0.0602)	(0.0792)	(0.0439)	(0.0401)	(0.0750)
Observations	2,644	3,571	4,316	5,911	8,228
Spatial Kernel Distance	5500km	5500km	5500km	5500km	5500km

*** p<0.01, ** p<0.05, * p<0.1
Conley spatially corrected standard errors in parenthesis.
All columns contain County/Município and year fixed effects

Table 8: Estimates by Quantiles of Population - USA

	(1) Unmasked NTL Average Radiance	(2) Unmasked NTL Average Radiance	(3) Unmasked NTL Average Radiance	(4) Unmasked NTL Average Radiance	(5) Unmasked NTL Average Radiance
GDP	0.269***	0.273***	0.304***	0.336***	0.446***
	(0.0463)	(0.0461)	(0.0562)	(0.0533)	(0.0608)
Pop	0.519***	0.504***	0.454***	0.374***	0.131*
	(0.0590)	(0.0600)	(0.0733)	(0.0703)	(0.0782)
Observations	13,218	11,688	9,090	6,264	4,286
Spatial Kernel Distance	5500km	5500km	5500km	5500km	5500km

*** p<0.01, ** p<0.05, * p<0.1
Conley spatially corrected standard errors in parenthesis.
All columns contain County/Município and year fixed effects

Table 9: Estimates by Quantiles of GDP - Brazil

	(1) Unmasked NTL Average Radiance	(2) Unmasked NTL Average Radiance	(3) Unmasked NTL Average Radiance	(4) Unmasked NTL Average Radiance	(5) Unmasked NTL Average Radiance
GDP	0.288***	0.229***	0.0893**	-0.217***	-0.338***
	(0.0678)	(0.0330)	(0.0413)	(0.0420)	(0.0407)
Pop	0.497***	0.558***	0.726***	1.090***	1.158***
	(0.0898)	(0.0418)	(0.0527)	(0.0552)	(0.0552)
Observations	11,200	10,275	9,525	7,932	5,615
Spatial Kernel Distance	5500km	5500km	5500km	5500km	5500km

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

Conley spatially corrected standard errors in parenthesis.

All columns contain County/Município and year fixed effects

Table 10: Estimates by Quantiles of Population - Brazil

References

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

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, 124(2):428–488.

Asher, S., Lunt, T., Matsuura, R., and Novosad, P. (2021). Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in india using the shrug open data platform. *The World Bank Economic Review*, 35(4):845–871.

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.

Banerjee, A. and Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in india. *American economic review*, 95(4):1190–1213.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). Roads, railroads, and decentralization of chinese cities. *Review of Economics and Statistics*, 99(3):435–448.

Berman, N., Couttenier, M., Rohner, D., and Thoenig, M. (2017). This mine is mine! how minerals fuel conflicts in africa. *American Economic Review*, 107(6):1564–1610.

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

Bluhm, R. and Krause, M. (2018). Top lights-bright cities and their contribution to economic development.

Bluhm, R. and McCord, G. C. (2022). What can we learn from nighttime lights for small geographies? measurement errors and heterogeneous elasticities. *Remote Sensing*, 14(5):1190.

Bruederle, A. and Hodler, R. (2018). Nighttime lights as a proxy for human development at the local level. *PloS one*, 13(9):e0202231.

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. (2015). A test of the new viirs lights data set: Population and economic output in africa. *Remote Sensing*, 7(4):4937–4947.

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.

Chen, X. and Nordhaus, W. D. (2019). Viirs nighttime lights in the estimation of cross-sectional and time-series gdp. *Remote Sensing*, 11(9):1057.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45.

Conley, T. G. and Molinari, F. (2007). Spatial correlation robust inference with errors in location or distance. *Journal of Econometrics*, 140(1):76–96.

Cook, C. J. and Shah, M. (2020). Aggregate effects from public works: Evidence from india. *The Review of Economics and Statistics*, pages 1–38.

Dalgaard, C.-J., Kaarsen, N., Olsson, O., and Selaya, P. (2018). Roman roads to prosperity: Persistence and non-persistence of public goods provision.

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

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., and Walker, M. W. (2019). General equilibrium effects of cash transfers: Experimental evidence from kenya. Technical report, National Bureau of Economic Research.

Elgin, C., M. A. K. F. O. and Yu., S. (2021). Growth and external debt. *CEPR Discussion Papers*, (16497).

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.

Elvidge, C. D., Baugh, K. E., Zhizhin, M., and Hsu, F.-C. (2013). Why viirs data are superior to dmsp for mapping nighttime lights. *Proceedings of the Asia-Pacific Advanced Network*, 35(0):62.

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.

Gennaioli, N., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics*, 128(1):105–164.

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

Gibson, J., Olivia, S., Boe-Gibson, G., and Li, C. (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics*, 149:102602.

Hao, R., Yu, D., Sun, Y., Cao, Q., Liu, Y., and Liu, Y. (2015). Integrating multiple source data to enhance variation and weaken the blooming effect of dmsp-ols light. *Remote Sensing*, 7(2):1422–1440.

Henderson, J. V., Squires, T., Storeygard, A., and Weil, D. (2018). The global distribution of economic activity: nature, history, and the role of trade. *The Quarterly Journal of Economics*, 133(1):357–406.

Henderson, J. V., Storeygard, A., and Deichmann, U. (2017). Has climate change driven urbanization in africa? *Journal of Development Economics*, 124:60–82.

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.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of Sciences*, 107(35):15367–15372.

Hu, Y. and Yao, J. (2021). Illuminating economic growth. *Journal of Econometrics*.

Huang, L. Y., Hsiang, S. M., and Gonzalez-Navarro, M. (2021). Using satellite imagery and deep learning to evaluate the impact of anti-poverty programs. Technical report, National Bureau of Economic Research.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794.

Jedwab, R., Kerby, E., and Moradi, A. (2017). History, path dependence and development: Evidence from colonial railways, settlers and cities in kenya. *The Economic Journal*, 127(603):1467–1494.

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

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.

Li, X., Xu, H., Chen, X., and Li, C. (2013). Potential of npp-viirs nighttime light imagery for modeling the regional economy of china. *Remote Sensing*, 5(6):3057–3081.

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).

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica*, 81(1):113–152.

Michalopoulos, S. and Papaioannou, E. (2014). National institutions and subnational development in africa. *The Quarterly journal of economics*, 129(1):151–213.

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

Ranjan, P. and Talathi, K. (2021). Impact of colonial institutions on economic growth and development in india: Evidence from night lights data.

Shi, K., Huang, C., Yu, B., Yin, B., Huang, Y., and Wu, J. (2014). Evaluation of npp-viirs night-time light composite data for extracting built-up urban areas. *Remote Sensing Letters*, 5(4):358–366.

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.

Tuttle, B. T., Anderson, S. J., Sutton, P. C., Elvidge, C. D., and Baugh, K. (2013). It used to be dark here. *Photogrammetric Engineering & Remote Sensing*, 79(3):287–297.

Online Appendix

This is the online appendix for *Planes, Trains and Automobiles: What Drives Nighttime Light?*

A Contemporaneous Daytime Images of Economic Development

(a) Seven Corners, VA

(b) College Park, MD

(c) Seven Corners, VA

(d) College Park, MD

Figure 1: Corresponding daytime images from Washington, DC Metropolitan Area; Source: Google Earth

(a) Memorial Bridge, DC/VA Border, October 2019

(b) Memorial Bridge, DC/VA Border, March 2020

(c) Memorial Bridge, DC/VA Border, June 2020

(d) Memorial Bridge, DC/VA Border, October 2020

Figure 2: Corresponding daytime images from Washington, DC Metropolitan Area; Source: Google Earth

B Masked, Unmasked and Gas-Flare-Removed Regressions

	USA			Brazil		
	(1) Sum of Avg. Yearly Radiance	(2) Sum of Avg. Yearly Radiance	(3) Sum of Avg. Yearly Radiance	(4) Sum of Avg. Yearly Radiance	(5) Sum of Avg. Yearly Radiance	(6) Sum of Avg. Yearly Radiance
GDP	0.00761**	0.0150**	0.636***	0.00254*	0.00255*	0.118***
	(0.00387)	(0.00592)	(0.0872)	(0.00149)	(0.00150)	(0.0331)
Population		-0.0222	-0.294***		-0.000549	0.256***
		(0.0208)	(0.0671)		(0.000378)	(0.0401)
Observations	24,670	24,670	24,670	44,547	44,547	44,547
Number of Counties/Municipios	3,086	3,086		5,569	5,569	
State-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Spatial Kernel Distance			5500km			5500km
Autoregressive shocks			Yes			Yes

Cols. 1-2, 4-5, cluster-robust standard errors in parentheses (county/município level)
Conley spatially corrected standard errors in parenthesis.
All columns contain county/município fixed effects
*** p<0.01, ** p<0.05, * p<0.1

Table 1: Comparison of Panel Data Regression via Traditional “Within” Procedure (Cols. 1 & 2, 4 & 5) versus Spatially-Corrected Estimates (Cols. 3 & 6)

	USA			Brazil		
	(1) Masked Sum of Avg. Yearly Radiance	(2) Unmasked Sum of Avg. Yearly Radiance	(3) Unmasked Sum of Avg. Yearly Radiance	(4) Masked Sum of Avg. Yearly Radiance	(5) Unmasked Sum of Avg. Yearly Radiance	(6) Unmasked Sum of Avg. Yearly Radiance
GDP	0.636***	0.636***	0.623***	0.104***	0.118***	0.116***
	(0.0868)	(0.0872)	(0.0878)	(0.0257)	(0.0331)	(0.0333)
Population	-0.292***	-0.294***	-0.279***	0.231***	0.256***	0.259***
	(0.0669)	(0.0671)	(0.0680)	(0.0506)	(0.0401)	(0.0404)
Observations	24,670	24,670	24,670	44,547	44,547	44,547
State-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Spatial Kernel Distance	5500km	5500km	5500km	5500km	5500km	5500km
Autoregressive shocks	Yes	Yes	Yes	Yes	Yes	Yes
Gas Flare Removal	No	No	Yes	No	No	Yes

*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Conley spatially corrected standard errors in parenthesis.
All columns contain county/município and year fixed effects.

Table 2: Estimates Masked Nighttime Lights Images vs. Unmasked Nighttime Lights Images

C Tests for Parameter Stability

	(1) Sum of Avg. Yearly Radiance	(2) Sum of Avg. Yearly Radiance	(3) Sum of Avg. Yearly Radiance	(4) Sum of Avg. Yearly Radiance
GDP	0.634***	0.634***	0.635***	0.630***
	(0.0909)	(0.0901)	(0.0899)	(0.0907)
Population	-0.290***	-0.288***	-0.289***	-0.289***
	(0.0696)	(0.0691)	(0.0691)	(0.0706)
Observations	21,585	21,585	21,585	21,587
State-Year FE	Yes	Yes	Yes	Yes
Spatial Kernel Distance	5500km	5500km	5500km	5500km
Autoregressive spatial shocks	Yes	Yes	Yes	Yes
Year Dropped	2012	2013	2014	2015

	(1) Sum of Avg. Yearly Radiance	(2) Sum of Avg. Yearly Radiance	(3) Sum of Avg. Yearly Radiance	(4) Sum of Avg. Yearly Radiance
GDP	0.632***	0.634***	0.634***	0.659***
	(0.0908)	(0.0913)	(0.0916)	(0.0894)
Population	-0.291***	-0.293***	-0.293***	-0.323***
	(0.0708)	(0.0712)	(0.0713)	(0.0667)
Observations	21,587	21,587	21,587	21,587
State-Year FE	Yes	Yes	Yes	Yes
Spatial Kernel Distance	5500km	5500km	5500km	5500km
Autoregressive spatial shocks	Yes	Yes	Yes	Yes
Year Dropped	2016	2017	2018	2019

*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Conley spatially corrected standard errors in parenthesis
All columns contain county/município and year fixed effects
All estimates use unmasked average radiance data

Table 3: Tests for Parameter Stability - USA

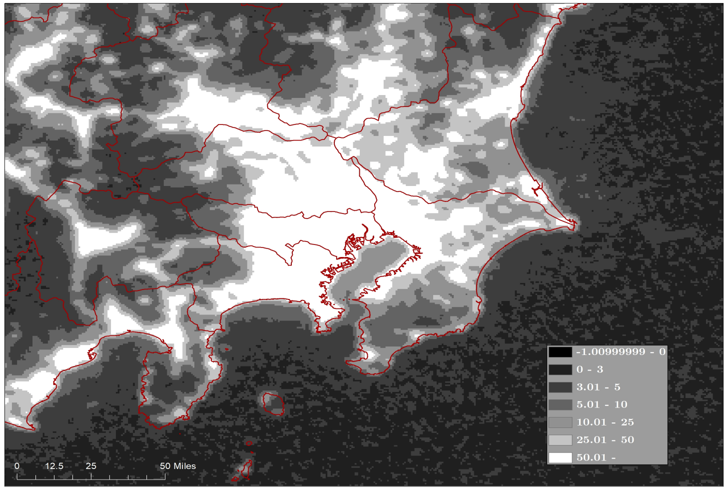
VARIABLES	(1) Sum of Avg. Yearly Radiance	(2) Sum of Avg. Yearly Radiance	(3) Sum of Avg. Yearly Radiance	(4) Sum of Avg. Yearly Radiance
GDP	0.122***	0.127***	0.122***	0.120***
	(0.0347)	(0.0344)	(0.0336)	(0.0338)
Population	0.251***	0.245***	0.252***	0.255***
	(0.0416)	(0.0415)	(0.0406)	(0.0408)
Observations	38,983	38,978	38,978	38,978
State-Year FE	Yes	Yes	Yes	Yes
Spatial Kernel Distance	5500km	5500km	5500km	5500km
Autoregressive spatial shocks	Yes	Yes	Yes	Yes
Year Dropped	2012	2013	2014	2015

VARIABLES	(1) Sum of Avg. Yearly Radiance	(2) Sum of Avg. Yearly Radiance	(3) Sum of Avg. Yearly Radiance	(4) Sum of Avg. Yearly Radiance
GDP	0.118***	0.120***	0.0956***	0.119***
	(0.0338)	(0.0349)	(0.0313)	(0.0354)
Population	0.256***	0.254***	0.285***	0.256***
	(0.0409)	(0.0423)	(0.0375)	(0.0430)
Observations	38,978	38,978	38,978	38,978
State-Year FE	Yes	Yes	Yes	Yes
Spatial Kernel Distance	5500km	5500km	5500km	5500km
Autoregressive spatial shocks	Yes	Yes	Yes	Yes
Year Dropped	2016	2017	2018	2019

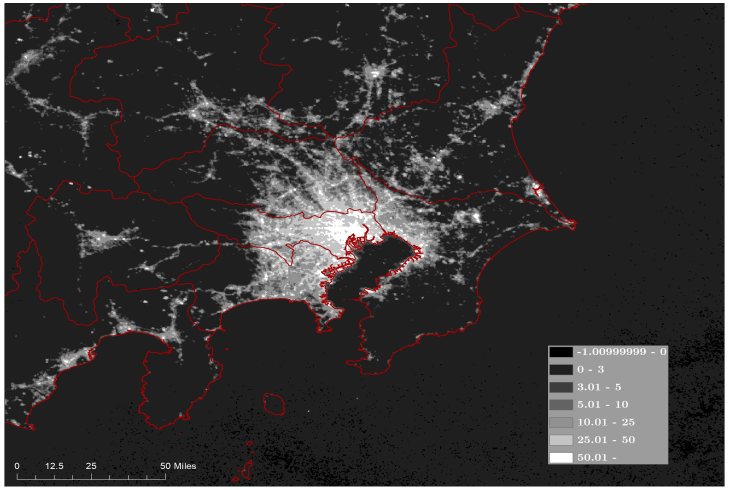
*** p<0.01, ** p<0.05, * p<0.1
Conley spatially corrected standard errors in parenthesis
All columns contain county/município and year fixed effects
All estimates use unmasked average radiance data

Table 4: Tests for Parameter Stability - Brazil

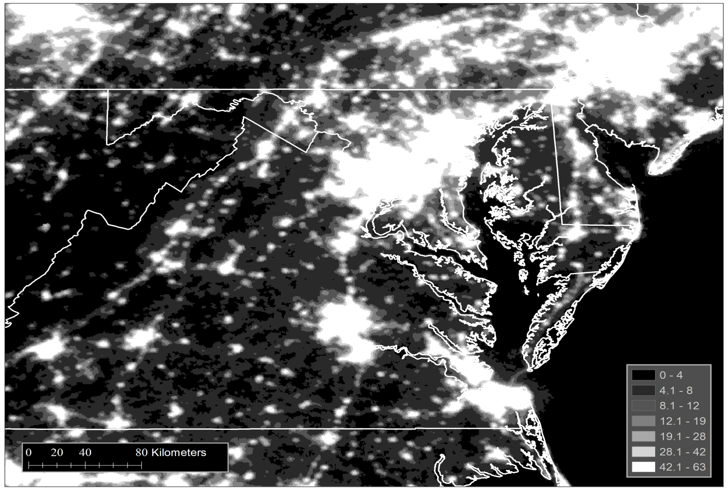
D VIIRS vs DMSP



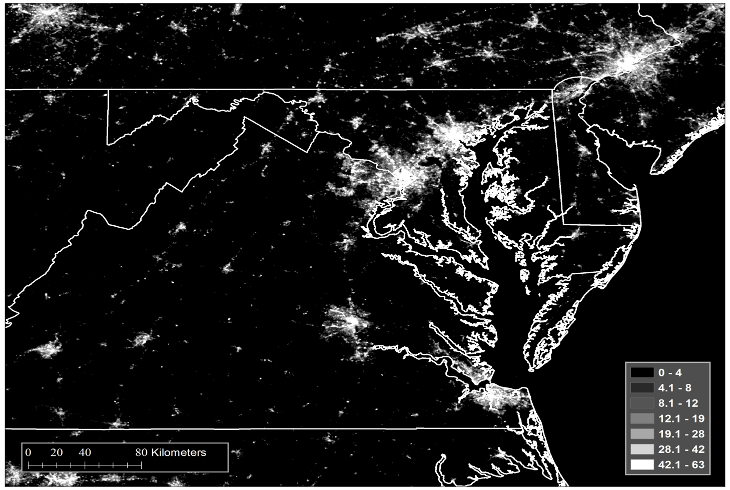
(a) Metropolitan Tokyo, DMSP 2013



(b) Metropolitan Tokyo, VIIRS 2013



(c) DC, Maryland and Virginia, DMSP 2013



(d) DC, Maryland and Virginia, VIIRS 2013

Figure 3: DMSP Nighttime lights(Older Generation, Top Panel) Contrasted with VIIRS Nighttime Lights (newer generation, bottom panel)

E Data Preparation

NTL VIIRS version 10, monthly cloud-free radiance with both masked and unmasked versions are the product that has been used. Masked products have been demonstrated in the literature to have a stronger relationship with economic activity [Gibson and Boe-Gibson, 2021].

> Prior to averaging, the DNB data is filtered to exclude data impacted by stray light, lightning, lunar illumination, and cloud-cover. Cloud-cover is determined using the VIIRS Cloud Mask product (VCM). In addition, data near the edges of the swath are not included in the composites (aggregation zones 29-32).
>
> Temporal averaging is done on a monthly and annual basis. The version 1 series of monthly composites has not been filtered to screen out lights from aurora, fires, boats, and other temporal lights. However, the annual composites have layers with additional separation, removing temporal lights and background (non-light) values.

Due to stray light, at higher latitudes no images are available for the summer months. This raises a question of how to average the nighttime light values across the year when some pixel-months have missing values. In order to resolve this issue the following formula has been used to create annual rasters at the pixel level for the entire globe from the monthly composites. All relevant data files are hosted online and fully available.

Take m as the month of the image and M is the number of monthly images available in a year. We can use a weighted sum of the monthly images to create an annualized raster.[1] This is done to compile the highest quality data up to the annual level. The method outlined here is helpful as it accounts for the fact that in 2012 only 9 months worth of monthly images are available to include when generating the annual composite. For this reason sensitivity analysis should be performed with 2012 data to ensure that error is not introduced by the lack of data from the first quarter of 2012.

$$\alpha_{py} = \frac{\sum_1^M (\frac{\alpha_{pm} \times \gamma_{pm}}{\gamma_{py}} + 1)}{M} \tag{1}$$

where:

$$\gamma_{py} = \sum_{m=1}^{M} \gamma_{pm}$$

To provide an example, if we have a pixel where γ_{pm} takes values $\{1, 23, 2, 0, 34, 3\}$ and α_p takes the values $\alpha_p \in \{15, 20, 14, 0, 45, 18\}$ and $\gamma_p = 63$

$$\alpha_{py} = \frac{\sum_1^6 (\frac{\alpha_{pm} \times \gamma_{pm}}{\gamma_{py}} + 1)}{6} \tag{2}$$

[1] A previous version of this paper summed all 12 monthly rasters and divided by 12 using the Arc GIS raster calculator. This is not appropriate as in the case of USA there are northern latitudes in the top half of the US particularly which do not have any satellite nighttime lights observations in the summer months due to glare.

$$\alpha_{py} = \frac{(\frac{(15\times1)}{63}+1)+(\frac{(20\times23)}{63}+1)+\cdots+(\frac{(18\times3)}{63}+1)}{6} \quad (3)$$

note that if the number of cloud free observations, γ_{pm}, is zero, then the product of $\alpha_{pm}\times\gamma_{pm} = 0$ this data point for the pixel is simply "dropped" from the average or included with a weight of zero. This procedure is of critical importance to account for stray light during the summer months at higher latitudes. This missing data affects a significant portion of the continental United States though Brazil is not affected. Previous authors have utilized a shorter time-span with prepared annual composite products.

The raster calculator function of ArcGIS was used to compute the annual weighted averages at the pixel level.

> In the monthly cloud-free DNB composites, there are many areas of the globe where it is impossible to get good quality data coverage for that month. This can be due to cloud-cover, especially in the tropical regions, or due to solar illumination, as happens toward the poles in their respective summer months. Therefore, it is imperative that users of these data utilize the cloud-free observations file and not assume a value of zero in the average radiance image means that no lights were observed.
>
> The version 1 global monthly VIIRS images are denoted with filenames containing "vcm" and these images exclude any data impacted by stray light.

E.1 Processing BLS GDP Data

The Bureau of Economic Analysis (BEA) who provide the GDP data at the county level, which are available from 2001-2019, combines some counties to simplify measurement and interpretation of their GDP estimates. Many counties in Virginia, as well as one in Hawaii are combined for the analysis. In order to do this the sum of light is calculated for all areas, and then simply added together for the combined statistical areas.

The US Census bureau provides annual estimates of county-level populations using American Community Survey 5-year data. In this manner all data have been adapted to match with the BLS GDP measurement areas.

E.2 Processing Gas Flare Masked Data

Data on the presence of gas flares was ascertained also from the Earth Observation Group at the Colorado School of Mines.

These data record the peak radiant emissions from flares based on thermal data analysis from the VIIRS instrument which detects and measures the volume of the radiant emissions from gas flares caused by the burning of natural gas using at night. The number of individual flare sites was estimated at 7,467 in 2012, for example [Elvidge et al., 2015].

In order to remove the gas flares from the nighttime lights data, I first imported the gas flare spatial points data into ArcGIS, and created a buffer zone of 1km around all gas flare locations.

This was done on a year-by-year basis so that gas flares from 2012 were only cleaned out of the 2012 night lights data and not the 2017 night lights data while 2017 gas flares, more specifically their buffered polygons, were removed from 2017 data using the erase feature of ArcGIS. I then re-calculated zonal statistics using the masked polygons to obtain the estimated sum of lights net the gas flaring regions.

References

Elvidge, C. D., Zhizhin, M., Baugh, K., Hsu, F.-C., and Ghosh, T. (2015). Methods for global survey of natural gas flaring from visible infrared imaging radiometer suite data. *Energies*, 9(1):14.

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