# 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 physical characteristics 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, accomplished 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

Social scientist's use of data on human activity extracted from nighttime satellite images has flourished over the past decade, catalyzed by groundbreaking papers from Henderson et al. (2012) and Chen and Nordhaus (2011). Nighttime lights images are free, high-quality, widely available and capable of providing insights that other data are unable to provide. The latest nighttime lights dataset, VIIRS nighttime lights, not yet widely adopted by the social sciences, provides images whose resolution is 45 times higher than the older generation (DMSP). Most significantly, nighttime lights provide high-resolution data on changes in human activity for most of the globe.

Though there are a few papers which document a correspondence between nighttime light and high-resolution GDP, previous nighttime lights papers have exclusively assessed the correspondence of GDP and light without controlling for the partial effect of changes in population on nighttime light.[1] If population changes directly influence the level of GDP in an area, and also directly influence the amount of light produced in the same area, then omitting population changes from the right hand side of the estimates is problematic due to classic omitted variables bias (Stock and Watson, 2020).[2]

If GDP and nighttime lights are more strongly related in one area, and population and nighttime light more strongly related in another area, that has implications for nighttime lights estimations that combine data across large areas, or from different countries. Estimating the strength 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 that undertakes 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

[1]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 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 and county and year fixed effects, rather than county-level fixed effects.

[2]Some authors have employed other strategies to address the influence of population on nighttime light including stratifying estimates across different population densities, or using population density at a particular point in time (Chen and Nordhaus, 2019; Gibson et al., 2021)

might indicate nighttime lights data is of little value in measuring such a project. Even with powerful fixed-effects procedures, combining areas in nighttime lights regressions may be problematic if nighttime lights measure population better in some places, and GDP better in others. How much are light changes driven by migration, for example, and how much by increases in output per worker? Again, without an econometric model that incorporates both elements simultaneously, we cannot know the answer. For these reasons, a primary objective of this paper will be to estimate the effect of GDP on nighttime lights, holding constant changes in the population, and estimate the effect of population on nighttime lights, holding GDP constant.

Despite the fact that we have no estimates of the direct effect of changes in the size of the population on nighttime light, researchers have been utilizing nighttime lights as a proxy for GDP at a high resolution (Hodler and Raschky, 2014; Kocornik-Mina et al., 2020). By incorporating a measure of population directly into the estimation, I attempt to evaluate these relationships at a high geospatial resolution, using advanced methods to account for the spatial relationships, heterogeneities, and nonlinear relationships that are present in the data. I believe this exercise will better equip future users of nighttime light, who will have a sharper understanding of whether nighttime light is capturing changes in economic output, or changes in the number of people.

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. An additional motivation of this paper is to compare the United States nighttime lights-GDP and nighttime lights-population estimates side-by-side with contemporaneous estimates from quality data, provided to the public by the Brazilian statistical agency. By using Brazil I am able to compare the USA to a large, middle-income economy with a greater amount of poverty and informality than the USA. Brazilian municípios and U.S. counties represent the second administrative level in their respective countries. They have some degree of overlap in economic and physical characteristics allowing for direct comparisons.[3]

Recent work on nighttime lights has proposed that there may be heterogenous effects with respect to the effect of GDP on nighttime lights depending on an area's characteristics. To the extent that county and município and state-year fixed-effects do not account for variation in the area's characteristics, I evaluate potential heterogeneity in the effect of GDP on nighttime lights by dividing the USA and Brazilian samples into quantiles corresponding to GDP and Population levels. I then re-estimate the econometric model for the partial effect of GDP on nighttime light and population on nighttime light, quantile by quantile, to compare results across quantiles for both Brazil and USA. I also divide the entire sample by percentiles of nighttime light separately for Brazil and the USA, and plot the estimated coefficients for each percentile for comparison.

With respect to econometrics methods applied to nighttime lights data, I argue that this paper makes several advancements by incorporating additional features novel to the GDP-nighttime lights literature into my econometric model. In a spatially dense area, it is likely that county and município-level economic shocks are correlated across space as well as time (Conley, 1999; Colella et al., 2019). If there are inter-dependencies between a county or município's unobservable characteristics and our variables of interest, GDP and population, in an estimation procedure this can affect the parameter and standard error estimates (Conley, 1999). In terms of nighttime lights-GDP estimates, the first new method that I employ is a procedure, developed by Conley (1999), to account for spatial relationships that other nighttime lights-

[3]Although other papers include additional countries, to the best of my knowledge there are no countries with population and GDP data available at a resolution sufficiently comparable to the second administrative levels of the USA and Brazil. Characteristics of municípios and counties are compared in table 3. 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). In theory this could 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. Source: https://www1.folha.uol.com.br/mercado/2020/02/informalidade-atinge-recorde-em-19-estados-e-no-df-diz-ibge.shtml

GDP estimates have avoided thus far. The Conley standard error technique has been leveraged in other applied economics work such as Hsiang (2010), Berman et al. (2017), Egger et al. (2019) and Sviatschi (2022). Another recent approach has demonstrated that, under certain conditions, arbitrary clustering of standard errors has been shown to be a better estimator of spatially structured standard errors (Colella et al., 2019). I apply this method alongside the Conley and Molinari (2007) spatial HAC corrected standard errors, and compare to a conventional fixed-effects estimator with White's heteroskedasticity-robust standard errors.

Another method unique to this paper is that, in addition to the normal controls for time-invariant, county-and-município-level, unobserved heterogeneity, the characteristics of the dataset permit controls for *state-year* unobserved heterogeneity. The inclusion of state×year dummies is similar in spirit to the use of country×year dummies in models of international trade. This accounts for state-specific annual shocks such as weather shocks, state elections, or localized inflation. That is, state×year fixed effects allow me to account for *time-variant*, state-year specific heterogeneity in my econometric estimates that would not otherwise be captured in a "normal" fixed-effects within-county or within-year transformation.[4] Consistent with findings in the nighttime lights-GDP literature, I also adopt and test the statistical significance of higher-order transformations of the independent variables (Hu and Yao, 2021).

Other authors have demonstrated that physical characteristics such as ports or links to navigable waterways correspond to nighttime lights activity (Henderson et al., 2018). Furthermore, the omission of public goods such as ports, border crossings, airports, roads, railways and the presence of navigable waterways from cross-sectional nighttime lights - GDP regressions can be problematic. These features may directly influence both economic growth (GDP) and nighttime light, or population and nighttime light. Therefore, in a separate specification I include controls for public goods, areas which,

[4]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

*a priori*, are known to be concentrated centers of economic activity. In many cases, the effect of time-invariant features is wiped out by panel data fixed-effects procedures. Using these data I collapse all GDP, nighttime light, and population observations to their county and município-level means for the years 2012-2020 in a between-county or between-município procedure. The between-county, between-município regression approach allows me to identify the contribution of public goods and physical characteristics to nighttime light despite their time-invariant nature. If these public goods and physical characteristics contribute significantly to nighttime light, controlling for population and GDP, that indeed confirms the importance of accounting for physical characteristics and infrastructure elements in estimating models that incorporate nighttime lights at a high geospatial resolution.

## 2 Literature Review

With respect to the growing 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.

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. I also attempt to assess the correspondence of physical characteristics with the production of nighttime light. Rather than including agricultural characteristics, elevation and latitude, I estimate the effects of time-invariant characteristics that are known to be concentrators of economic activity: ports, border crossings, airports, highways, railways and navigable waterways.

Mellander et al. (2015), using VIIRS nighttime lights data, 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. This finding, using high-resolution VIIRS data and high-quality Swedish administrative data, supports the need to establish the difference between the effects of population changes from the effect of changes in GDP holding population constant. An important limitation of the econometric analysis in Mellander et al. (2015) is that, though they do use the latest VIIRS data, the authors use cross-sectional rather than panel data. This makes it difficult for the authors to include controls for unobserved local public goods and other factors that might influence population, GDP or nighttime light.

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*. The effect of GDP on night-time 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. In the context of Levin and Zhang (2017), country fixed effects mask important sub-national heterogeneity, one of the motivations for the inclusion of county-level and município-level fixed effects in my econometric model.

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 disaggregate their 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. The econometric challenge of including population changes is therefore unresolved by their estimation procedure. Unfortunately, the DMSP data the authors use are known to display "blooming" or "bleeding" effects, where light measured at the sensor spreads from one pixel into the next (Hao et al., 2015). VIIRS imagery suffers no blooming effects, and provides resolution sufficient to address the question of the strength of the GDP-nighttime lights and population-nighttime lights relationship at the within-county and within-município level. Unlike the authors in Gibson and Boe-Gibson (2021), I am able to take advantage of a greater number of annual estimates of VIIRS nighttime light by combining monthly VIIRS nighttime lights images using a weighted average, with each monthly image weighted by the number of cloud-free images captured at the sensor during that particular month.[5]

Bluhm and McCord (2022) use the older DMSP nighttime lights and test the elasticity of nighttime light with respect to GDP at the county level for the USA and município level for Brazil. The authors recognize the shortcomings of the DMSP data, such as top-coding, bottom-coding, and a lack of an automatic gain sensor. The newer VIIRS data used in this paper are no longer constrained by the limitations outlined in Bluhm and McCord (2022) with respect to nighttime lights measurement. Similar to this paper, Bluhm and McCord (2022) also take the approach of estimating the nighttime lights production function, though they use light per area ($\frac{NTL}{AREA}$) as the dependent variable, where this paper uses the sum of all county or município light pixels.

In their estimates with the older satellite data, the Bluhm and McCord (2022) incorporate county and município fixed effects that would control for unobserved, time-invariant county and município-level characteristics. In terms of time fixed effects, however, the highest granularity the authors offer is country-year fixed effects. The setup the authors use accounts for any common shocks at the *country-level* in a particular year, for 5 countries. *State-level*, time-variant heterogeneity and shocks such as weather events, political shocks, or localized inflation would remain present in the data.

The authors include a measure of population, though they use only the initial population at the start of their data records, interacted with the level of GDP. They make no justification for including an interaction term while omitting the population variable measured in levels. In the case of my analysis, I directly incorporate a measurement of changes in the population for both the USA and Brazil, as well as the interaction

[5]Annual VIIRS nighttime lights images are only available to the public for specific years. I was able to extend the VIIRS data to include all the years since the satellite came online. Further details on how the nighttime lights images were compiled is available in the online data appendix.

of population and GDP to evaluate amplified or dampened effects at higher levels of population and GDP or lower levels of population and GDP.

Using the breadth of their sample, the authors in Bluhm and McCord (2022) also test for heterogeneity in terms of the effect of GDP on nighttime lights. The authors propose the strength of this relationship might vary across the spectrum of GDP. In this paper the estimates of the effect of GDP on nighttime light are separated by quantiles of GDP, and by quantiles of population. It may also be the case that at higher levels of population, the relationship between GDP and nighttime light or population and nighttime light could be stronger.

Three other recent studies evaluate the utility of VIIRS data for economic analysis at high geospatial resolutions. Chen and Nordhaus (2015) combine DMSP nighttime lights and data on output and population from Kenya. The author's data are much briefer than those included in this analysis, and there are no estimates for partial correlations that include population. 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 the presence of local public goods. Public goods and certain physical characteristics have been shown to be significant contributing factors to nighttime lights (Henderson et al., 2018). Hu and Yao (2021) find evidence of nonlinearities in the GDP-lights relationship at the country level. The findings from Hu and Yao (2021) are directly incorporated and tested in my analysis.

I view this paper as being related most closely to or a successor of Henderson et al. (2012), though the paper is closely related to Mellander et al. (2015), Bluhm and McCord (2022), and Hu and Yao (2021). My analysis takes advantage the substantial increase in the resolution at which nighttime lights images are available, as well as the lack of measurement issues related to DMSP nighttime lights measurement. Henderson et al. (2012) and Bluhm and McCord (2022) used 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. None of the above studies includes controls for changes in the level of the population.

Though Gibson and Boe-Gibson (2021) analyzes the same data at the county-level in the USA, the authors do not include direct measures of population in their estimates and they do not draw comparison across countries with similar characteristics. The analysis in Gibson and Boe-Gibson (2021) avoids addressing the potential for complications due to spatial dependencies in the data. I am able to leverage more VIIRS annual images than Gibson and Boe-Gibson (2021) by compiling monthly VIIRS images into annual averages, weighted by the number of monthly cloud-free observations captured by the sensor. This paper thus fills a needed gap to test whether light is a reasonable proxy for GDP in higher-income areas, or if in lower-income areas changes in nighttime lights may better capture changes in population than changes in GDP. It is my hope that this paper can set the stage and provide support for many more nighttime lights papers to come.

# 3 Methodology

To derive 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.

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, it is therefore better practice to use a potentially noisy variable as the dependent variable (Stock and Watson, 2020). In that way any measurement error may have less influence on the estimation of our coefficients of interest. 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 address the potential for spatially-correlated economic shocks using the procedures developed by Conley (1999), Conley and Molinari (2007). This method uses a non-parametric bootstrap estimator of the covariance to account for the underlying spatial structure of the data. I take advantage of the flexibility this estimator offers to allow for 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. Under 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. I also apply a newer procedure based on research by Colella et al. (2019) and an associated statistical package, that creates arbitrary standard error clusters. The arbitrary clustered standard errors have been shown to be slightly more accurate in some circumstances than the Conley and Molinari (2007) HAC standard errors.

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, $\gamma_c$ are the county/município fixed effects and $\varepsilon_{ct}$ is an idiosyncratic error term. In addition to the county-level controls for county-and-município-level unobserved heterogeneity, I include state-year fixed effects, $\psi_{st}$, which control for *time-variant*, unobserved, state-year specific economic shocks such as weather shocks, price shocks, political elections or other state-level unobserved time-variant economic volatility. Though computationally expensive, I argue these results allow the most robust and precise estimates of the effect of GDP on lights. I use the total 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 GDP 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. Consistent with Bluhm and McCord (2022), 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 2012-2020. 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 (2012-2020) or in many cases these features take many years to prepare and construct. The estimated equation using the between estimator is:

$$\widehat{NTL_c} = \beta[\widehat{GDP_c}] + \alpha[\widehat{POP_c}] + \phi_1[\text{Port}_c] + \phi_2[\text{PrimaryRoad}_c] + \phi_3[\text{Airport}_c] + \\ \phi_4[\text{Rail}_c] + \phi_5[\text{BorderPoint}_c] + \phi_6[\text{Water}_c] + \psi_s + \varepsilon_c \tag{3}$$

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

# 4 Data

I contrast data and estimates from the United States and Brazil, two countries which have some similar characteristics and some differences.[6] Both Brazil and the United States feature diverse geographic characteristics including mountains, lakes, rivers and coastlines as well as vast networks of infrastructure.[7] Though municípios are, on average, smaller than counties, there is significant overlap between município size and county size. 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.

[6]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.

[7]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. 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.

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 Brazilian municípios. 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, at the time of writing, we have no GDP estimates at the município level past 2020 for Brazil.

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

Table 1: Data Availability

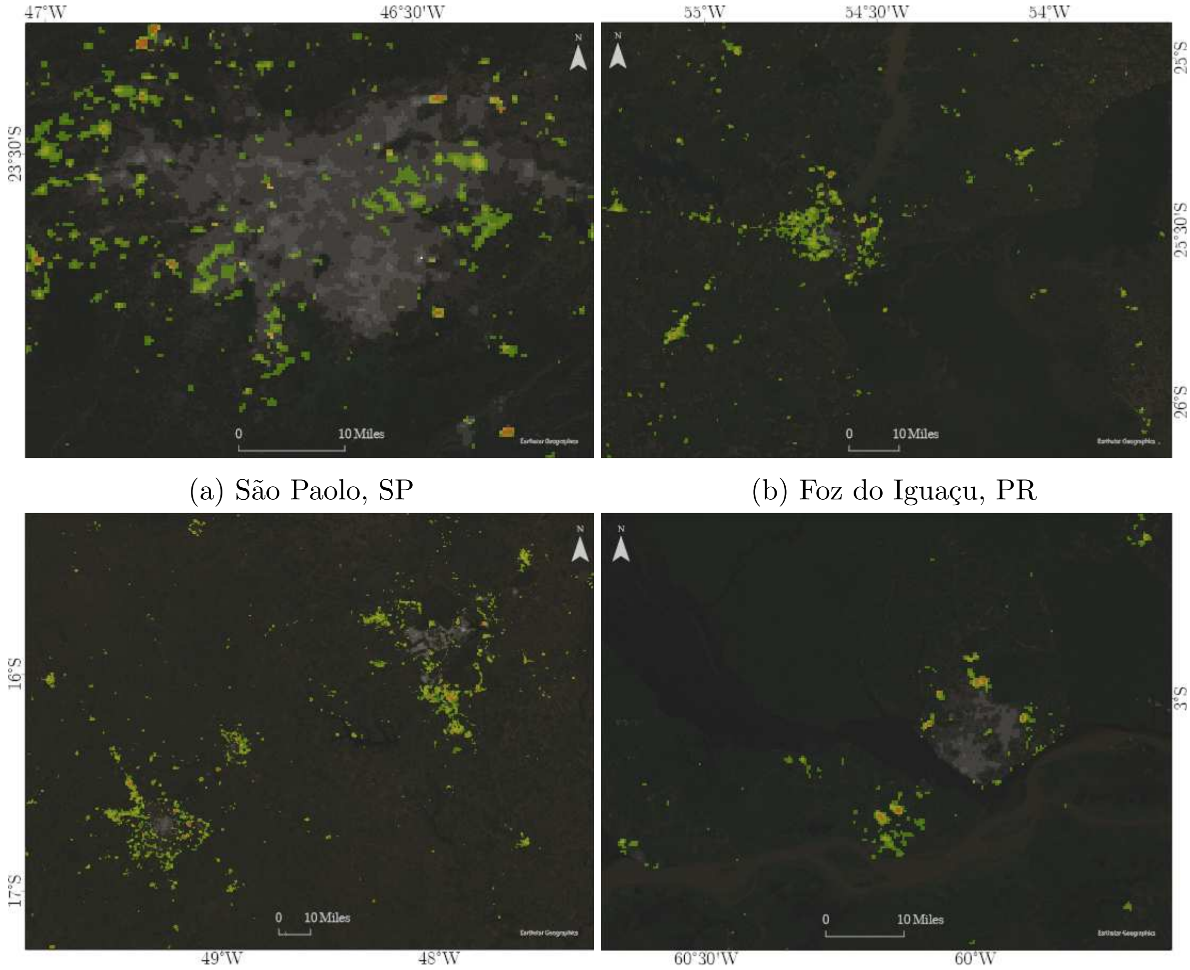
## 4.1 BLS/IBGE GDP Data

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. 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).[8] 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

[8]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).

compiled from governmental and other administrative data sources, similar to the U.S.A. GDP estimates.[9]

[9]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)**

## 4.2 ACS/IBGE County-Level and Município-level Population Data

Population estimates come from the American Community Survey (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 município population estimates also come from the IBGE. The município-level estimates are based on population growth projections, derived from the Brazilian population census in 2000 and 2010. The estimates are adjusted to match the growth rates of the states in which the municípios exist (IBGE, 2019).

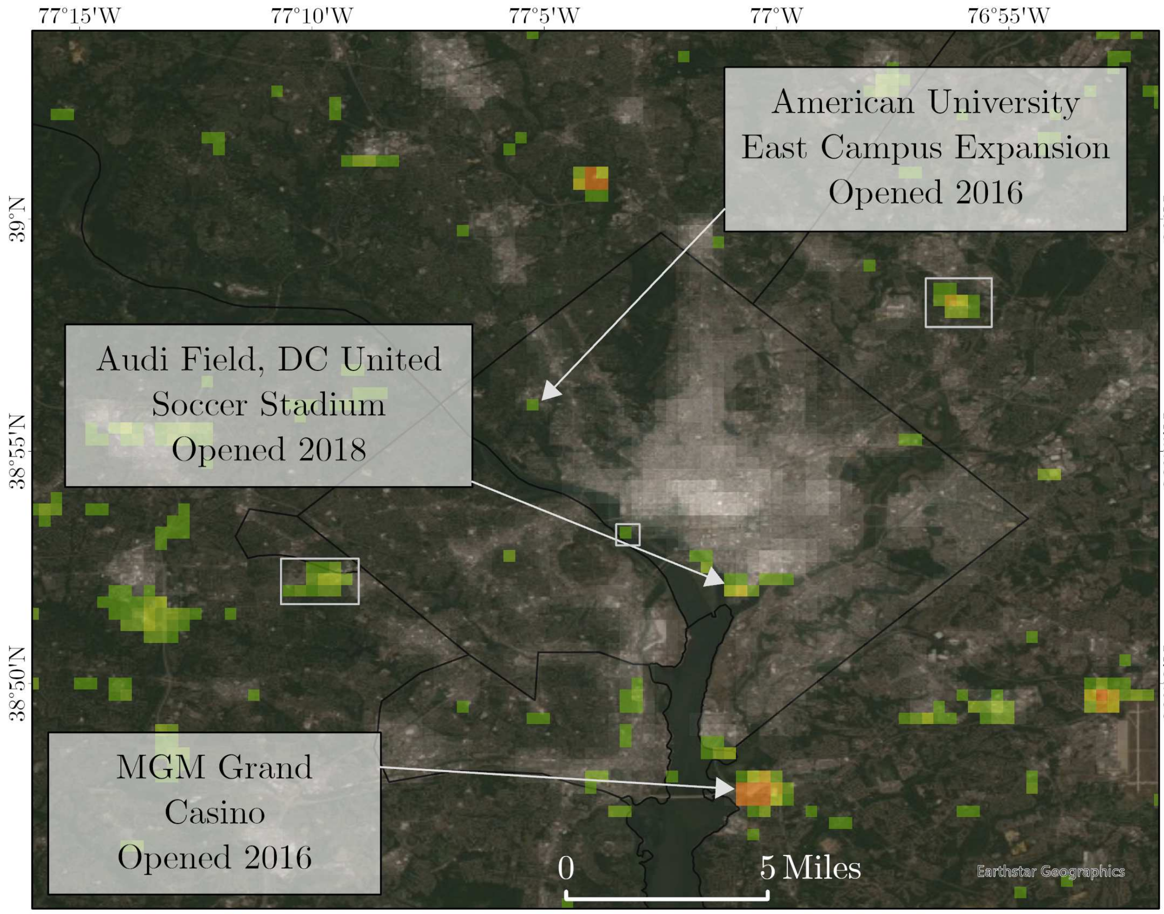



Figure 2: Night-time growth in Washington, DC 2012-19. Green = small change, Red = large change. Contemporaneous daytime imagery of economic development in the boxed areas is shown in the appendix. Layers: Basemap: Open Street Map, CC License; Night-time Lights Annual Image (2019) **(TO BE PRINTED IN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

## 4.3 VIIRS Night-time Lights Data

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.[10] 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

[10]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.

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.[11] 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).[12]

VIIRS 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 sensor incorporates automatic adjustments that can better capture much lower and higher levels of light than the previous generation (Elvidge et al., 2017).[16] The new VIIRS images are available on a daily frequency or in monthly

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

[12] VIIRS stands for Visible Infrared Imaging Radiometry Suite

[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 means 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.

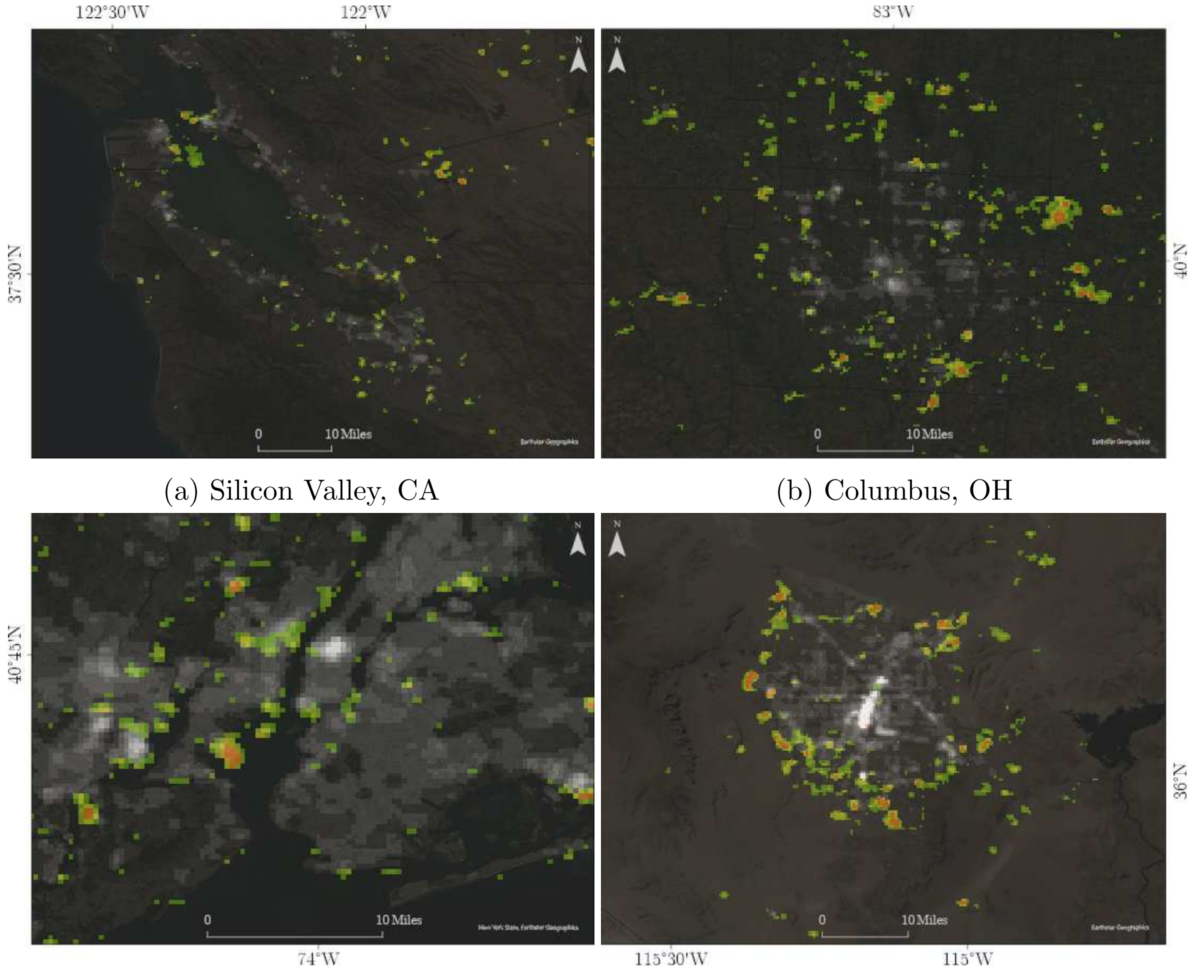
[16] The Suomi-NPP satellite flies over the earth around 1:30am and 1:30pm local time each day and captures images using the spectroradiometer, a device similar to the capture device in a digital camera (Carlowicz, 2012). Raw data from the sensor are then processed to remove non-human

composite forms as well as some yearly composite images (Carlowicz, 2012). [17] For the data estimated here I utilize the V1 monthly VIIRS nighttime images compiled into annual composite images using a weighted sum.[18]

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

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




(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)**

Looking directly at images of long run changes in nighttime light can illustrate the capacity of nighttime light to identify localized economic growth. 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.

## 4.4 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.

[19] https://ourairports.com/
[20] https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2020&layergroup=Roads
[21] https://inde.gov.br/

| | (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 |
|---|---|---|---|---|---|
| GDP | 0.0158*** | 0.0159*** | 0.00883 | 0.635*** | 2.117*** |
| | (0.00576) | (0.00576) | (0.0350) | (0.0834) | (0.197) |
| Population | | -0.0122 | 0.344*** | -0.293*** | -0.960*** |
| | | (0.0153) | (0.0488) | (0.0636) | (0.258) |
| $GDP^2$ | | | 0.00101 | | -0.0502*** |
| | | | (0.000825) | | (0.0180) |
| $Pop^2$ | | | -0.0169*** | | 0.0987*** |
| | | | (0.00320) | | (0.0191) |
| GDP×Pop | | | -0.00195 | | -0.0665* |
| | | | (0.00194) | | (0.0350) |
| Observations | 27,788 | 27,788 | 27,788 | 27,788 | 27,788 |
| State-Year FE | Yes | Yes | Yes | Yes | Yes |
| County/Municipio FE | Yes | Yes | Yes | Yes | Yes |
| Spatial Kernel Distance | - | - | - | 5500km | 5500km |
| Autoregressive spatial shocks | No | No | No | Yes | Yes |

Cols. 1,2 - Cluster-Robust standard errors in parentheses
Cols. 3,4 - Conley spatially corrected standard errors in parenthesis.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: Nighttime Lights Regressions with State-Year Dummies - USA

| | (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 |
|---|---|---|---|---|---|
| GDP | 0.00298* | 0.00300* | 0.0330*** | 0.124*** | 1.195*** |
| | (0.00154) | (0.00154) | (0.00896) | (0.0313) | (0.118) |
| Population | | -0.00146** | 0.0481*** | 0.249*** | 0.602*** |
| | | (0.000573) | (0.0168) | (0.0376) | (0.129) |
| $GDP^2$ | | | -3.50e-07 | | 0.00513 |
| | | | (0.000471) | | (0.00684) |
| $Pop^2$ | | | -0.000622 | | 0.0687*** |
| | | | (0.00103) | | (0.0132) |
| GDP×Pop | | | -0.00324*** | | -0.127*** |
| | | | (0.00122) | | (0.0167) |
| Observations | 50,125 | 50,125 | 50,125 | 50,125 | 50,125 |
| State-Year FE | Yes | Yes | Yes | Yes | Yes |
| Município FE | Yes | Yes | Yes | Yes | Yes |
| Spatial Kernel Distance | - | - | - | 5500km | 5500km |
| Autoregressive spatial shocks | No | No | No | Yes | Yes |

Cols. 1,2 - Cluster-Robust standard errors in parentheses
Cols. 3,4 - Conley spatially corrected standard errors in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Nighttime Lights Regressions with State-Year Dummies - USA

# 5 Results

## 5.1 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 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, $\beta$, are consistent with the intuition that a greater amount of output corresponds to a greater amount of light. The estimated effect of population on nighttime light is negative, which goes against the intuition that each individual consumes a certain amount of light. 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 (about \$1.2 million) of GDP. 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

[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.

population distribution of the USA and climbs to turn positive around the middle of the distribution at 10 log points, which corresponds to around 22,000 people.

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, $\beta$ & $\alpha$, I next decompose these estimates for different slices of the distribution in the following section.

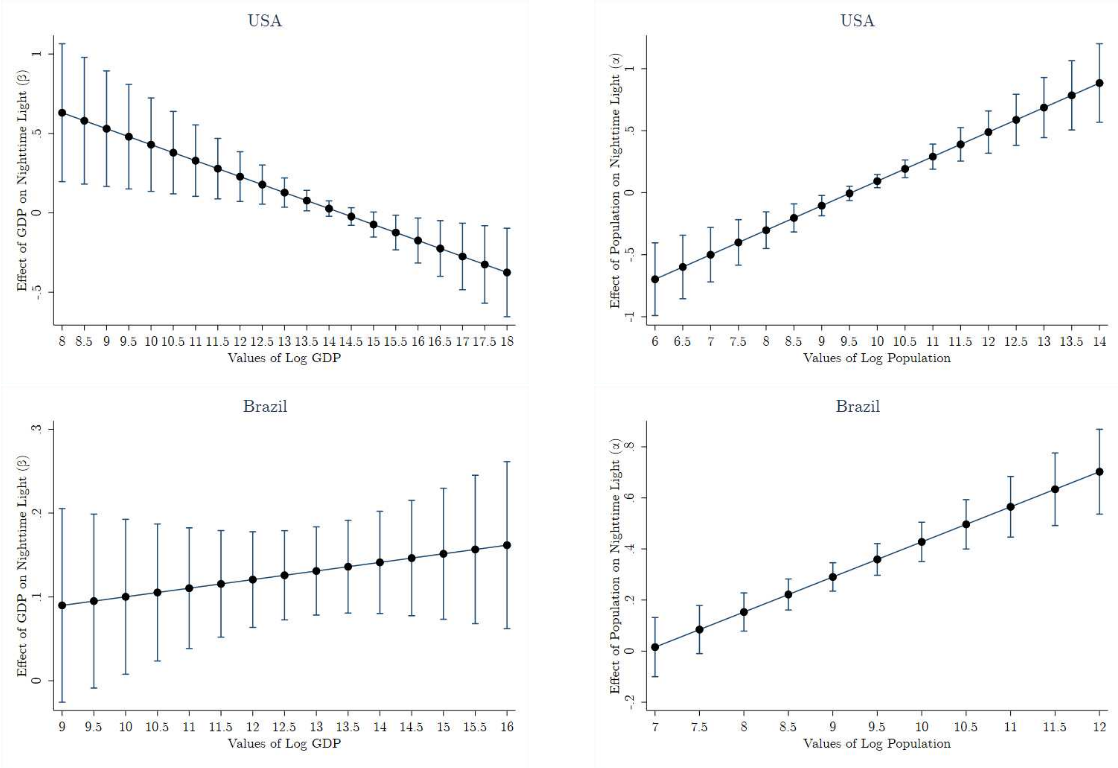



Figure 4: Marginal Effects

## 5.2 Regressions by Quantiles

The following analysis of the effect of population and GDP, $\alpha$ and $\beta$, on nighttime light divides the sample into quantiles of GDP, population and area. In each case the quantiles are standardized, and estimates can therefore be compared from the lowest-income Brazilian municípios with the poorest USA counties. Table 8 in the appendix 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.

### 5.2.1 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

| | | GDP Quantile (1) | GDP Quantile (2) | GDP Quantile (3) | GDP Quantile (4) | GDP Quantile (5) |
|---|---|---|---|---|---|---|
| USA | GDP - $\beta$ | 0.782 | 1.232 | 1.16 | 0.925 | 0.836 |
| | Pop - $\alpha$ | 0.134 | -0.613 | -0.573 | -0.319 | -0.292 |
| BRA | GDP - $\beta$ | 0.269 | 0.273 | 0.304 | 0.336 | 0.446 |
| | Pop - $\alpha$ | 0.519 | 0.504 | 0.454 | 0.374 | 0.131 |

Complete regression tables included in appendix tables 8-10.
Spatially adjusted standard errors used in all estimates.
All estimates contain County/Município and year fixed effects.

Table 4: Estimated Coefficients by Quantiles of GDP

the individual quantiles. In the USA, across all quantiles the GDP effect dominates the estimated population effect. Interestingly, in 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 a barrier for researchers seeking to use nighttime lights data within the same country, across different sections of the distribution of GDP.

### 5.2.2 Quantiles of Population

Table 5 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, $\alpha$ appears to be larger than the effect of GDP on nighttime light. In Brazil we see a similar pattern. Even in the lowest quantiles, the effect of population on nighttime light, $\alpha$, is estimated to be smaller than the effect of increases in GDP. Moving up

| | | Population Quantile (1) | Population Quantile (2) | Population Quantile (3) | Population Quantile (4) | Population Quantile (5) |
|---|---|---|---|---|---|---|
| USA | GDP - $\beta$ | 0.617 | 0.599 | 0.442 | 0.317 | 0.275 |
| | Pop - $\alpha$ | 0.298 | 0.207 | 0.358 | 0.476 | 0.439 |
| BRA | GDP - $\beta$ | 0.288 | 0.229 | 0.0893 | -0.217 | -0.338 |
| | Pop - $\alpha$ | 0.497 | 0.558 | 0.726 | 1.09 | 1.158 |

Complete regression tables 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.

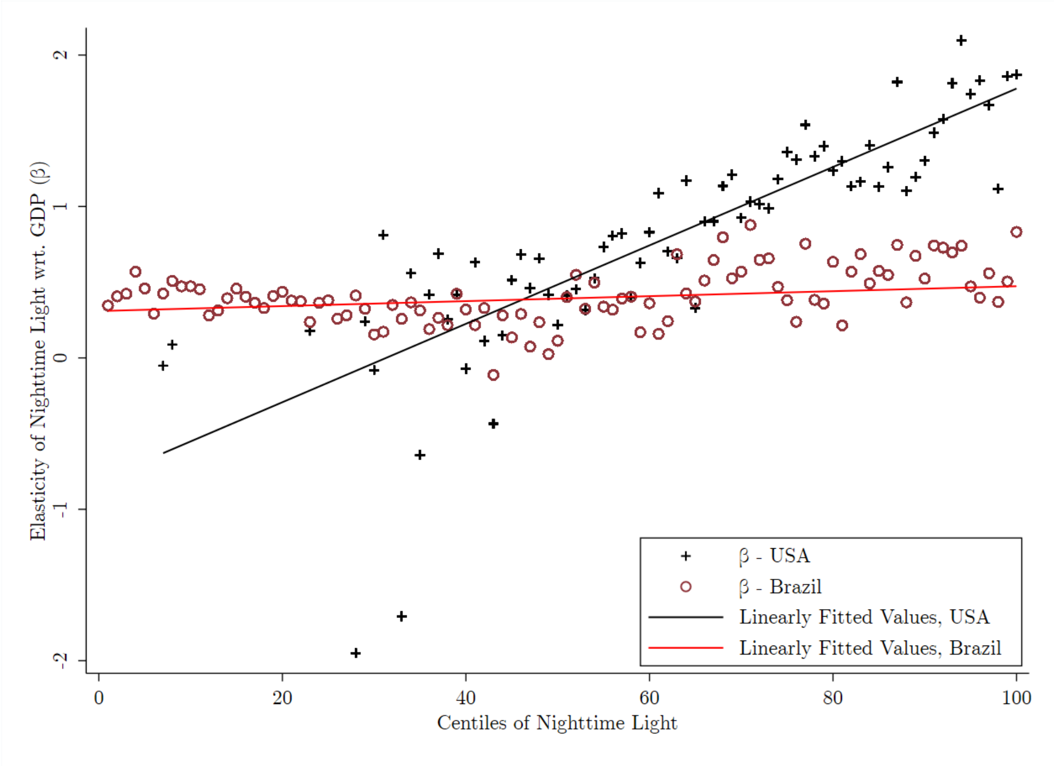
Table 5: Estimated Coefficients by Quantiles of Population

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 greater the influence of population changes on 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.

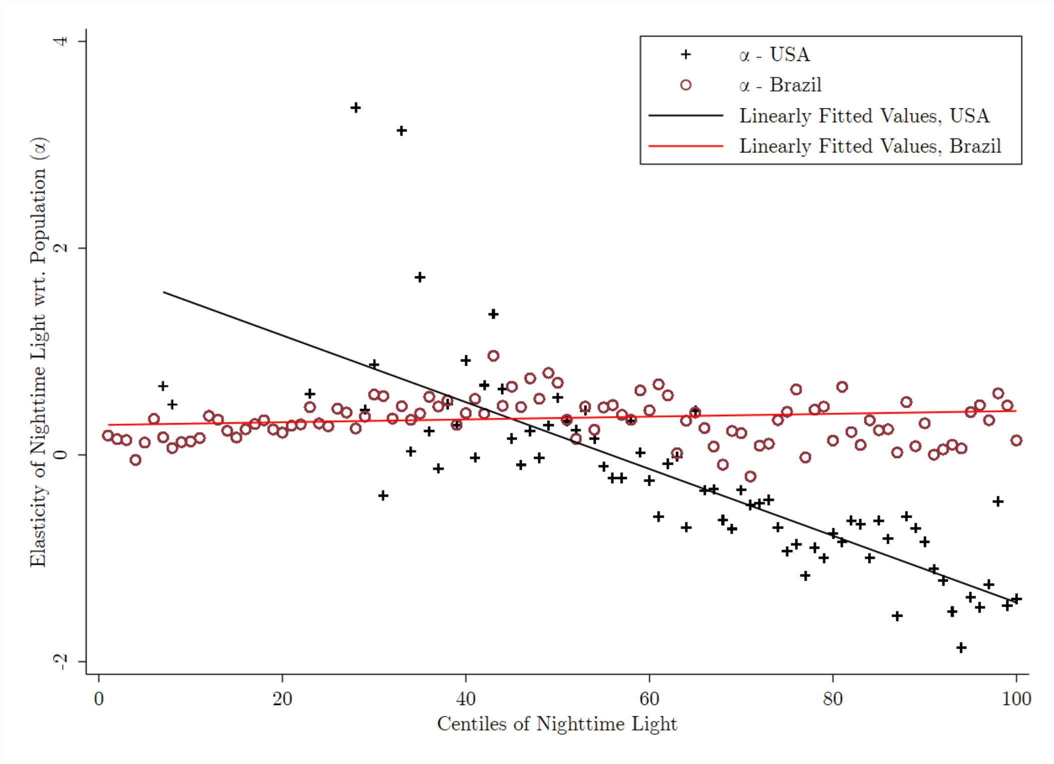
## 5.3 Regressions by Centile

Figure 5 examines the estimates of the effects population and GDP, $\alpha$ and $\beta$, on the sum of light. I divided the distribution of nighttime light into hundredths (percentiles or centiles), and estimated the model 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, $\beta$, 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 light, $\alpha$, 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.




(a) Effect of GDP on Nighttime Light




(b) Effect of Population on Nighttime Light

Figure 5

## 5.4 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 2012-2020. 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$\times$year dummy variables to control for time-variant state-specific heterogeneity. One exception is the GDP$\times$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 $\phi_1$ to $\phi_6$ 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.

# 6 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, $\beta$, 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 $\alpha$ 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.

# 7 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) Unmasked NTL Average Radiance | (2) Unmasked NTL Average Radiance | (3) Unmasked NTL Average Radiance | (4) 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 6: 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 |

$^{\dagger}$nW/cm2/sr

Table 7: Descriptive Statistics for All Regression Variables

# Appendix

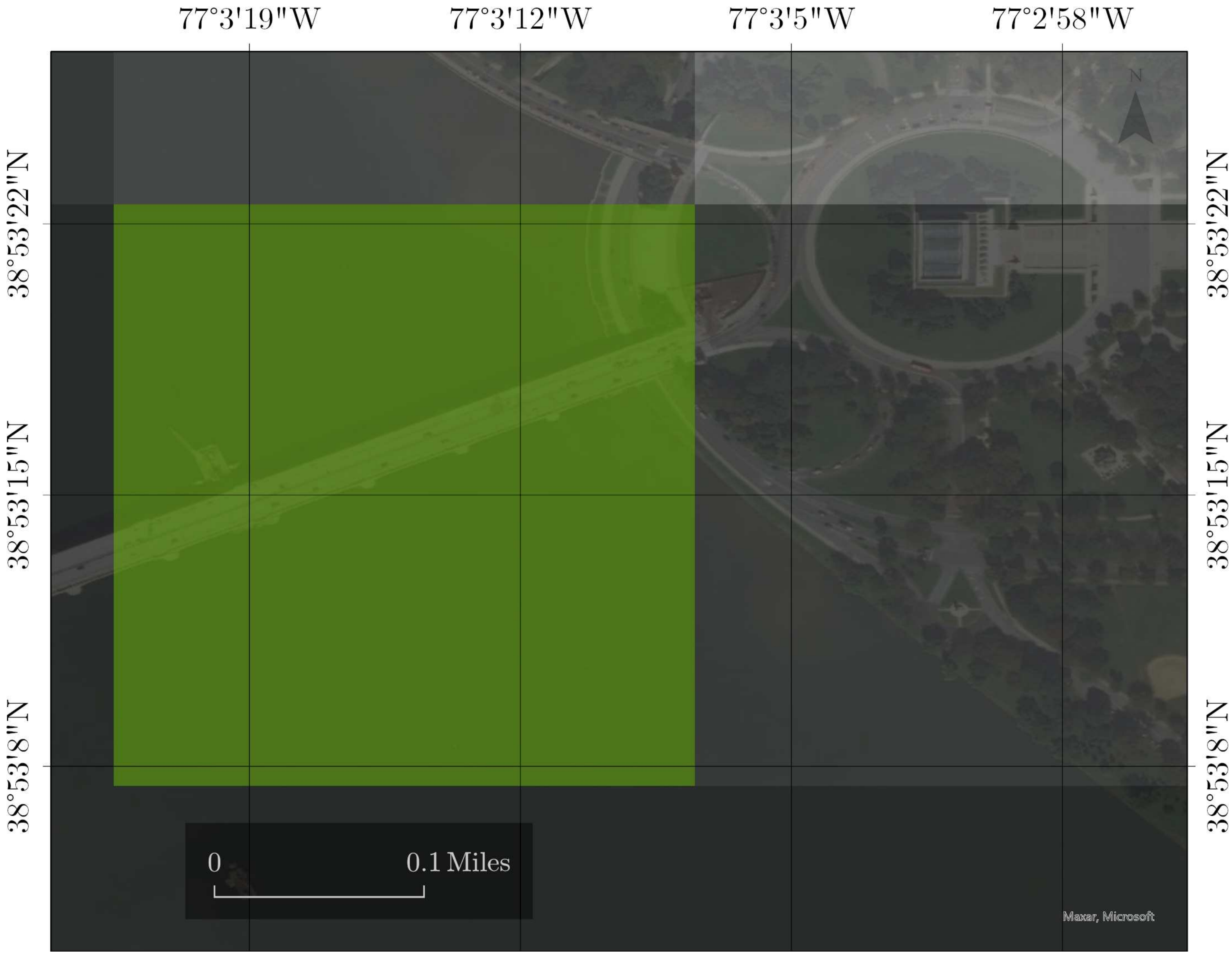



Figure 6: Night-time growth Memorial Bridge. Layers: Basemap: Open Street Map, CC License; Night-time Lights Annual Image (2019); Green = small change, Red = large change **(TO BE PRINTED IN COLOR)**

| | 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 8: Characteristics of Counties  Municípios by Quantiles of Size

| | (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 9: 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 10: 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 11: 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 12: 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.

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.

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 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.

Colella, F., Lalive, R., Sakalli, S. O., and Thoenig, M. (2019). Inference with arbitrary clustering," iza discussion papers 12584. *Institute of Labor Economics (IZA)*.

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.

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.

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 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.

IBGE (2019). Estimativas da população residente para os municípios e para as unidades da federação brasileiros com data de referência em 1º de julho.

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

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

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

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.

Stock, J. H. and Watson, M. W. (2020). Introduction to econometrics 4th ed.

Sviatschi, M. M. (2022). Making a narco: Childhood exposure to illegal labor markets and criminal life paths. *Econometrica*, 90(4):1835–1878.

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.