Article

Nighttime Lights and County-Level Economic Activity in the United States: 2001 to 2019

John Gibson * and Geua Boe-Gibson

Department of Economics, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand; geua.boegibson@waikato.ac.nz
* Correspondence: jkgibson@waikato.ac.nz; Tel.: +64-7-838-4289

Abstract: Nighttime lights (NTL) are a popular type of data for evaluating economic performance of regions and economic impacts of various shocks and interventions. Several validation studies use traditional statistics on economic activity like national or regional gross domestic product (GDP) as a benchmark to evaluate the usefulness of NTL data. Many of these studies rely on dated and imprecise Defense Meteorological Satellite Program (DMSP) data and use aggregated units such as nation-states or the first sub-national level. However, applied researchers who draw support from validation studies to justify their use of NTL data as a proxy for economic activity increasingly focus on smaller and lower level spatial units. This study uses a 2001–19 time-series of GDP for over 3100 U.S. counties as a benchmark to examine the performance of the recently released version 2 VIIRS nighttime lights (V.2 VNL) products as proxies for local economic activity. Contrasts were made between cross-sectional predictions for GDP differences between areas and time-series predictions of GDP changes within areas. Disaggregated GDP data for various industries were used to examine the types of economic activity best proxied by NTL data. Comparisons were also made with the predictive performance of earlier NTL data products and at different levels of spatial aggregation.

Keywords: VIIRS; DMSP; GDP; nighttime lights; cross-sectional; time-series; economic statistics

1. Introduction

Satellites have been observing the Earth at night for over 50 years, but it is especially since the digital archive of nighttime lights (NTL) was established in 1992 by the National Oceanic and Atmospheric Administration (NOAA) that researchers have found an ever-growing set of use for these data. Several key early studies by non-economists showed that NTL data from the Defense Meteorological Satellite Program (DMSP) could be used to estimate sub-national indicators of economic activity and per capita incomes [1–5]. Potential advantages of these NTL-based estimates, compared to traditional economic activity statistics like national or regional gross domestic product (GDP), are timelines, lower cost, comparability between countries irrespective of statistical capacity, and availability for spatial units below the level at which GDP data are reported.

In the last decade, economists have also begun using NTL data. Widely cited early studies from two different research teams noted that DMSP data are noisy, but in a wide range of contexts [6,7], or alternatively, just in data-poor environments [8,9], DMSP data could add value to conventional economic statistics. In contrast to earlier studies focused particularly on comparing regions, a theme in recent studies by economists is using NTL data to track fluctuations in local economic activity in response to various shocks such as disasters [10–12], or certain policy interventions [13,14]. This use of NTL as a proxy for changes in local economic activity, plus ongoing cross-sectional use as a proxy for variation in economic performance, raises the question of how predictive NTL data are for studying differences in economic activity between areas and the temporal changes in activity within areas.

Several validation studies have considered this question by using GDP data as a benchmark for assessing predictive performance of NTL data. An early and widely cited study used national level DMSP and GDP data for 188 countries from 1992 to 2008 [7], while a similar study used these data for 1500 regions (mostly at first sub-national level) from 82 countries from 1992 to 2009 [15]. However, applied researchers who draw support from validation studies to justify their use of NTL data as an economic activity proxy have increasingly focused on smaller and lower level spatial units [16]. Several studies have used DMSP data at the third sub-national level, which includes counties, sub-districts, and NUTS3 regions [10,17–20], with some studies for even lower level spatial units such as villages [14], micro-grids [21], and even pixel-level [11,22]. A mismatch exists between the spatial level of validation studies and the spatial level of applied studies that use NTL data to proxy for economic activity matters because flaws in DMSP data such as spatial imprecision and blurring [23,24] make the predictive performance far worse for lower level spatial units such as the third sub-national level than for more aggregated units such as the national or first sub-national level [25].

The extant validation studies are mainly for older NTL data products such as DMSP. Some comparisons between GDP and version 1 NTL annual composites from the Visible Infrared Imaging Radiometer Suite (VIIRS) have been made [26], but these products are only for 2015 and 2016. To date, no validation studies have used version 2 VIIRS annual composites (V.2 VNL), which have recently been released [27]. To help close this gap in the literature, this study used the 2001–19 time-series of GDP for over 3100 U.S. counties as a benchmark to examine the usefulness of three NTL data sources, DMSP, V.1 VNL, and V.2 VNL as proxies for local economic activity. We included data from the 2014–18 extension of DMSP based on pre-dawn readings (compared to the early evening readings for DMSP prior to 2014). We also used the V.2 VNL data with two other samples, a cross-country dataset, so that results could be compared with earlier validation studies [7] and state-level U.S. data to examine the aggregation effects. Our panel data estimation framework helps to contrast cross-sectional predictive performance for differences between areas with performance for a time-series of changes within areas. A further contribution is to use GDP for various industries to see what economic activities are best proxied by NTL data. The industry-level results and related split-sample results based on agriculture's contribution to GDP and on population density provide a basis to consider how our findings may apply to other settings where the economic structure differs from that of the United States.

2. Materials and Methods

2.1. Related Literature on NTL Validation Studies

In the current context, validation studies have attempted to estimate the nature of the relationship between NTL data and traditional economic activity data for places with trustworthy data. These studies provide a basis for using NTL data as a proxy in other times and places where traditional data such as GDP are either absent or not trusted. The errors in GDP data should be independent of errors in NTL data, so some studies have noted an optimal indicator of true economic activity would weight a mixture of the two measures [7–9]. Studies using this framework have put some weight on DMSP data for examining cross-sectional differences in places where the GDP data have low reliability, but note that without further refinement of the NTL data, they are "not a reliable proxy for time-series measures of output growth" [9] (p. 241). A far lower predictive ability for time-series changes, even if DMSP data are good predictors of cross-sectional differences in economic performance, also holds at very local (third sub-national) levels in a developing country setting [28].

The VNL data from VIIRS are a refinement over DMSP data, in terms of spatial precision and temporal consistency [23], so the question of whether these data are a reliable proxy for measuring changes in economic activity has been examined, albeit within the limits of the short time-series for V.1 VNL annual composites. The V.1 VNL data predict over 70% of variation in U.S. state-level GDP (and over 85% of variation in GDP for

metropolitan areas), but predict less than 4% of variation in annual rates of change in GDP [26]. Direct comparisons of VIIRS and DMSP have been limited because the V.1 VNL annual composites are only for 2015–16 [29] and the popular DMSP stable lights time-series [30] ends in 2013 (data from the DMSP 2014–18 extension are yet to be used). To deal with this issue, annual NTL estimates for 2013 from VIIRS monthly data are constructed by various researchers, usually with masking procedures to remove outliers in the monthly data, and these VIIRS annual estimates better predict in cross-sections of GDP than DMSP data [25,31–33].

While several studies have noted that DMSP data are noisy measures of true luminosity, the nature of the measurement error has rarely been examined. A study at the second sub-national (NUTS2) level for Europe found mean-reversion, where errors in DMSP data negatively correlate with true values [33]. Unlike random errors that do not bias regression coefficients if NTL data are the left-hand side variable and attenuate coefficients in proportion to the reliability ratio if they are the right-hand side variable [34], mean-reverting errors in a left-hand side variable cause bias and in a right-hand side variable may overstate coefficients rather than attenuate them [35–37]. A decomposition using DMSP data adjusted for top-coding [38] found that most of the spatially mean-reverting errors were still present, implying that the blurring of the DMSP images [24] is the more important source of error in DMSP data [33].

A consequence of mean-reverting errors is understated inequality between places as NTL estimates revert toward their mean. Some studies have considered inequality as an aspect of economic performance by using DMSP data as a proxy in places that lack timely or fine resolution sub-national GDP data [39,40]. However, validation studies show that DMSP data understate spatial inequality, especially in urban and high density areas, with this pattern holding across developed and developing regions of the world [25,33].

Validation studies have also examined the types of economic activity (and hence, the type of places, given different patterns of specialization) for which NTL data are a poor proxy. The GDP-luminosity relationship (using DMSP data from 1992 to 2009) is positive for countries with agricultural shares of GDP below 20%, but negative elsewhere [41]. The weaker relationship with agricultural sector activity is also seen at the third sub-national level in China in the DMSP data, while the V.1 VNL data (annual estimates from masked monthly records) are unrelated to primary sector GDP [25]. If NTL data poorly capture agricultural activity, it may help explain why NTL data are a weaker proxy for economic activity in low density areas [42], given the predominance of agriculture in such places.

2.2. Data and Methods

We used four data sources to test the relationships between night lights and county-level and state-level GDP. The first was real GDP in chained 2012 dollars, from the U.S. Bureau of Economic Analysis (BEA). The annual estimates are provided separately for each county for the 2001 to 2019 period, except in Alaska, where the BEA combines some census areas in their reporting, for example, in Hawaii, where they combine Maui and Kalawao counties, and in Virginia, where there are 23 BEA-created combination areas where one or two independent cities with 1980 populations of less than 100,000 are combined with an adjacent county. The dissolve function in ArcGIS was used to modify a county-level shapefile, so that it matched these combination areas. There were n = 3109 counties and combination areas (we refer to all of these as county-level units) with data available in each year.

The second data source was four annual products for the 2014 to 2019 period from the version 2 VIIRS nighttime lights (V.2 VNL) annual composites [27]. We used the average radiance, median radiance, and the masked variants of these two data products, summing the radiance by county-level unit in each year. While the V.2 VNL annual composites are also available for 2012 and 2013 (as they are built from monthly data available since April 2012), the values for those two years are yet to have a stray light adjustment. With the northerly latitude of much of the U.S., stray light can affect the images on many nights.

This reduces comparability with the time-series from 2014 onwards, which is based on stray light corrected data, so we did not use the 2012 and 2013 V.2 VNL data.

The V.2 VNL are produced from monthly cloud-free radiance averages, with initial filtering to remove extraneous features such as fires and aurora before the resulting rough annual composites are subjected to outlier removal procedures. To isolate the background from lit grid cells, a data range threshold is set from 3×3 blocks of grid cells where the threshold is based on a multiyear maximum median and a multiyear percent cloud-cover grid [27]. In other words, there is a single data range threshold across all the years in the series, in contrast to the year-specific thresholds that were used for the version 1 VIIRS annual composites [29]. The data are in units of nano Watts per square centimeter per steradian ($nW/cm^2/sr$) reported on a 15 arc-second output grid.

The third data source was the version 1 VIIRS nighttime lights (V.1 VNL) annual composites for 2015 and 2016 [29]; the only two years for which this product is available. We used the stray light corrected version (vcmsl) of these annual composites, with the outliers removed and background set to zero (ormntl). The average annual radiances from each of the 15 arc-second output pixels were summed to county-level totals.

The fourth data source was annual composites from the Defense Meteorological Satellite Program (DMSP) satellites F14, F15, F16, and F18. These composites provide an average digital number (DN) for each 30 arc-second output pixel, where DN values are 6-bit digital numbers that range from 0–63, with higher numbers indicating greater brightness. Ephemeral lights such as from fires and gas flares are removed from the annual composites, and the original processing by NOAA scientists also excluded (at pixel level) images for any nights affected by clouds, moonlight, sunlight, and other glare. The usual stable lights product has a time-series that ended in 2013 [30], with two satellites providing data for each year up to 2007, so there are 20 satellite-years available over the 2001 to 2013 period.

The DMSP satellites have an unstable orbit, tending to observe Earth earlier as they age. For example, a satellite tracking mission (see: http://www.remss.com/support/crossing-times/ accessed on 6 August 2019) shows equator crossing times for F18 of 8 pm in 2013, but 6 pm by 2018. Thus, what starts out as a Day–Night observation becomes Dawn–Dusk observation. The Earth Observation Group at the Colorado School of Mines has exploited this feature to extend the time-series of DMSP stable lights annual composites by using pre-dawn data from satellite F15 for 2014 to 2018. Lights observed in the early hours of the morning are more likely to be from public infrastructure (e.g., street lights) than from private consumption and production activities, so the extended DMSP stable lights series may not be consistent with the earlier DMSP data, and we treated them as a separate source of information on NTL. For both sets of DMSP data, we used the sum of the DN values within a county-level unit.

Our main parameter of interest was the elasticity of GDP with respect to night lights, as estimated from the following regression:

$$\ln(real\ GDP)_{it} = \alpha + \beta \ln(sum\ of\ lights)_{it} + \mu_i + \varphi_t + \varepsilon_{it} \tag{1}$$

where the *i* indexes the cross-sectional units (county-level units in most cases but we also estimated Equation (1) with country and state-level data); the *t* indexes years; the μ_i are fixed effects for each cross-sectional unit; the φ_t are the fixed effects for each year; and ε_{it} is the disturbance term. The fixed effects let us control for time-invariant features of each cross-sectional unit, and spatially-invariant features of each time period. One could allow time effects to vary across space at some more aggregated level (e.g., at state level if there are county fixed effects), but the setup we used is the traditional one in economics studies using night lights data. The elasticity is a unit-free measure showing by what percentage the left-hand side variable changes for each percentage change in the right-hand side variable. Thus, the fact that the V.1 and V.2 VNL data are measured in $nW/cm^2/sr$ while the DMSP data are in DN values does not affect the estimation of the elasticity.

The specification of Equation (1) with NTL data on the right-hand side does not imply that lights cause GDP (as any causation would go the other way) and instead, it has a

predictive interpretation. The typical situation where NTL data are used as a proxy for local economic activity is because traditional measures like GDP are either unavailable or are considered untrustworthy. Thus, it is important to learn from settings like the U.S., where the GDP data are both available and trustworthy, about how closely NTL data correlate with GDP data, in order to see if the NTL data are an adequate proxy measure.

For example, many studies use NTL data to estimate impacts of a shock such as a natural disaster [10–12], which affects some cross-sectional units but not others, and occurs in some time periods but not others. The validity of using NTL data to estimate the impacts on local economic activity of such shocks (or more generally, of 'treatments') depends on the product of two relationships: $(\partial GDP / \partial lights) \cdot (\partial lights / \partial treatment)$. In the settings of interest, typically the $\partial GDP / \partial lights$ relationship is not estimated because there are no GDP data (as any available and trustworthy GDP data would already be used for the evaluation). Instead, the validation studies from elsewhere provide evidence on the $\partial GDP / \partial lights$ term that is needed for interpreting estimates of the impact of the treatment on night lights as estimates of the impact of the treatment on local economic activity. In other words, if relationships between changes in GDP and changes in NTL data are very weak, then it is hard to see how estimates of the $(\partial lights / \partial treatment)$ effect are informative about how the shock impacts on economic activity and performance.

To provide a basis to interpret results of Equation (1), we considered two widely cited studies (with 1850 and 650 *Google Scholar* citations as of May 2021) that have reported estimates of Equation (1). With 17 years of DMSP data for 188 countries, the elasticity is about 0.3 (long differences give a similar value) [7]. With 18 years of DMSP data for 1500 regions (typically at the first sub-national level) from 82 countries, an even larger elasticity of about 0.4 was reported [15].

The Equation (1) specification is known as a 'fixed effects' or 'within' estimator, as the variation that allows β to be estimated comes from time-series changes for each cross-sectional unit. In other words, Equation (1) lets one see how changes in annual GDP vary with changes in NTL data. An alternative estimator that uses the same panel data is the 'between' estimator, where averages over time for each cross-sectional unit are used in the regression (e.g., the average GDP of a county from 2014 to 2019 is regressed on the average sum of lights in the county over the same period). The between estimator allows for examination of cross-sectional GDP differences between areas while the within estimator allows for time-series predictions of GDP changes within areas. We report the results for both estimators. The NTL data have been used in various studies in both contexts; to proxy for economic performance in cross-sectional studies such as when long-run impacts of historical factors are considered [43], and in studies focused on fluctuations in economic activity because the intervention or shock that they study occurs in the sample period [12,44].

3. Results

3.1. Country-Level Results

We started with country-level results for a comparison to a key study that found a GDP-lights elasticity of 0.3 using the within estimator and DMSP data [7]. In the first two columns of Table 1, we show the results for all countries with data on real GDP in local currency units from 2014 to 2019 in the *World Development Indicators* [series NY.GDP.MKTP.KN].

The estimated GDP-lights elasticity was only 0.015 if the V.2 VNL average radiance product was used, while it was six-times larger, at 0.094, if the masked average was used. It seems that background noise and ephemeral sources of light in the unmasked data may attenuate within estimates of the elasticity. However, even after removing noise by masking, the elasticity was less than 0.1, which was far smaller than the earlier estimate of 0.3 with DMSP data. Moreover, omitting countries not in the sample of the widely cited Henderson et al. study [7] slightly lowered the estimated elasticity to 0.085 (column (3)). The other change in specification for results in the last two columns of Table 1 was to divide the sum of radiance by country area to match the way NTL data were

used in the Henderson et al. study, and to add a quadratic term for the model reported in column (4); the squared term is statistically insignificant ($p = 0.95$) and the double logarithmic specification seems appropriate.

Table 1. Within estimator results for GDP-lights elasticities using V.2 VNL data: country-level 2014 to 2019.

Independent Variables and Summary Stats	All Countries with Data		Henderson et al. (2012) Specification	
	Mean Radiance	**Masked Mean**	**Masked Mean/Area**	**Quadratic Model**
ln(lights)	0.015 **	0.094 **	0.085 **	0.084 **
	(0.007)	(0.038)	(0.037)	(0.040)
R-squared (within)	0.012	0.076	0.074	0.074

Notes: Based on a panel of 203 countries (1192 observations) in columns 1 and 2, with ln(lights) based on the sum of radiance by year and country. Columns 3 and 4 are based on 181 countries (1072 observations) using lights per square mile in column 3 (and a quadratic of this, with an unreported squared term, in column 4). Models include year and country fixed effects. Standard errors in parentheses clustered at country level, ** $p < 0.05$.

The results in Table 1 suggest that findings from earlier periods using DMSP data may not apply in more recent periods with VIIRS NTL data. However, there are at least two issues with this evidence. First, applied studies are increasingly focused on lower level spatial units, so country-level results may provide less guidance than in the past when the NTL data were used with more aggregated spatial units. The second and more concerning issue is that country level GDP data are of widely varying reliability and so they may not provide the consistent benchmark given by sub-national GDP data for the United States.

3.2. Results at County and State Level

The results of using four V.2 VNL products (average radiance, median radiance, masked average radiance, and masked median radiance) for a panel of 3109 county-level units observed each year from 2014 to 2019 are reported in Table 2. The top panel has the "within" estimator results, based on time-series variation, and the bottom panel has "between" estimator results, based on differences in average economic performance in the cross-section. Unlike the country-level results in Table 1, which are subject to wide variation in statistical capacity between countries that make some GDP data more trustworthy than others, we considered that county-level GDP data produced by the BEA will provide a consistent level of reliability over time and space. Consequently, differences in the lights-GDP relationships are interpreted in terms of potential measurement error features of the NTL data, rather than reflecting possible errors in the GDP data that may vary with either spatial scale or types of economic activity.

Table 2. Relationships between VIIRS V.2 NTL and county GDP: within and between estimator results.

Independent Variables and Summary Statistics	V.2 VNL Annual Data Product			
	Average Radiance	**Median Radiance**	**Masked Average Radiance**	**Masked Median Radiance**
Within-estimator, for annual GDP changes within each county				
ln(sum of lights)	0.021 *	0.004	0.118 ***	0.131 ***
	(0.013)	(0.011)	(0.019)	(0.022)
Year fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
R-squared (Within)	0.002	0.000	0.029	0.030
Between-estimator, for average GDP differences between counties				
ln(sum of lights)	1.261 ***	1.270 ***	1.049 ***	1.045 ***
	(0.015)	(0.015)	(0.007)	(0.008)
R-squared (Between)	0.706	0.709	0.863	0.861

Notes: Based on a strongly balanced panel of 3109 county-level units, observed each year from 2014 to 2019, giving N = 18,654 observations. Standard errors in parentheses (clustered at county level for the within-estimator results), * $p < 0.10$, *** $p < 0.01$.

The masked products were better predictors of time-series changes in GDP and cross-sectional differences in GDP than were the unmasked data products. The within-estimator R^2 values (which are always very low across all NTL data products, levels of spatial aggregation, types of economic activity, and time periods used in this study) were three points higher when using the masked data products. The between estimator R^2 values were 15 points higher (at 0.86 vs. 0.71) when using the masked VNL data products rather than their unmasked counterparts. Prior studies have shown that NTL data are more powerful cross-sectional predictors of differences in GDP (and other economic activity indicators) between areas than they are predictors of time-series changes [26,28,45]. This pattern also holds for the masked V.2 VNL data, where the R^2 values for the between estimator in the cross-section were almost 30 times as high as for the within-estimator of the time-series changes.

The GDP-lights elasticity was almost zero if using the within estimator with unmasked data products, and was 0.12 (0.13) when the masked average (median) was used. The masking procedure was designed to remove background noise and ephemeral sources of light [27]. To the extent that such noise is not auto-correlated across years, the usual pattern of random measurement error in a right-hand side variable, causing attenuation of the regression coefficient on that variable [34], seems to occur here, given that the estimated elasticity rises when masking is used to remove this noise from the data.

With this attenuation bias pattern in mind, it may seem puzzling that the between estimator results showed a larger GDP-lights elasticity (at 1.26 rather than 1.05) when the unmasked data products were used. Although not reported in Section 3.1, a similar pattern showed up in the country-level results, where the between-estimator gave a GDP-lights elasticity of 0.96 with the unmasked data and of 0.86 with the masked data (and the difference was statistically significant at $p < 0.02$). A potential explanation lies in the impact of non-random, and specifically mean-reverting, measurement errors. The unmasked data included occurrences of apparent light (either ephemeral or noise) outside of usually lit areas. After averaging across years, the apparent radiance of these unlit areas was raised and so the apparent luminosity of these areas became closer to the mean. With this mean-reverting error, when NTL data are on the right-hand side of a regression, the coefficients can be exaggerated, as seen in the first two columns of between estimator results in Table 2. Once this noise is removed, the results in the last two columns in the lower panel of Table 2 suggest that, on average, a county where the sum of NTL is ten percent higher than for another county will have a real GDP that is 10.5 percent higher.

The results in Table 2 are atypical of studies that relate NTL data to GDP data. While there are some county-level results for China [25], the validation studies with GDP data as a benchmark are mostly for spatially aggregated data at the national or first subnational level, even as applied studies increasingly use NTL data locally [45]. It is therefore of interest to see how the results for estimating Equation (1) change when the GDP and NTL data are at the state-level. This spatial aggregation suppresses much of the variation in the fluctuations; for example, the coefficient of variation for annual changes in log GDP, which is what the within-estimator is based on, has a value at the state level that is just one-sixth of the value at the county level. There is less suppression of variation for the between-estimator based on the averages over 2014–19, with the state-level coefficient of variation being one-half the county-level coefficient of variation.

An important change with state-level data is that there is less gain from masking to remove noise when using the within estimator; the top panel of Table 3 shows that the unmasked V.2 VNL data gives elasticities for changes in state-level GDP with respect to changes in state-level NTL of about 0.05, compared to 0.04 with the masked products (and these coefficients are surrounded by standard errors of about 0.03, so we cannot reject the hypothesis that the four sets of within-estimator elasticities in Table 3 are all the same). Unlike with the county-level data, predictive accuracy for annual changes in log GDP was not any higher when using the VNL masked data, and actually fell slightly from 0.05 to 0.02 (for the average radiances).

Table 3. Relationships between VIIRS V.2 NTL and state-level GDP: within and between estimator results.

Independent Variables and Summary Statistics	V.2 VNL Annual Data Product			
	Average Radiance	Median Radiance	Masked Average Radiance	Masked Median Radiance
Within-estimator, for annual GDP changes within each state				
ln(sum of lights)	0.050 **	0.047	0.043	0.037
	(0.025)	(0.032)	(0.038)	(0.031)
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
R-squared (Within)	0.053	0.040	0.021	0.013
Between-estimator, for average GDP differences between states				
ln(sum of lights)	0.598 ***	0.591 ***	0.840 ***	0.838 ***
	(0.116)	(0.114)	(0.083)	(0.079)
R-squared (Between)	0.351	0.355	0.679	0.699

Notes: Based on a strongly balanced panel of 51 state-level units (treating the District of Columbia as equivalent to a state), observed each year from 2014 to 2019, giving N = 306 observations. Standard errors are in parentheses (clustered at state level for the within-estimator results), ** $p < 0.05$, *** $p < 0.01$.

One interpretation of the fact that using masked data has little effect on the within estimator at the state level, unlike at the county-level, is that noise in estimates of annual changes in lights may cancel out as data are spatially aggregated to the state-level (noting also that there is less variability in annual GDP changes at state level than at county level). However, with even further aggregation to the country level in Table 1, using the masked data again seemed to matter (although discussion of the country-level relationships must be tempered by the fact that the GDP data across countries are likely to be a less consistent benchmark than are the sub-national data for the U.S. given the variation in statistical capacity between countries). The issue of how relationships between changes in NTL data and changes in GDP vary by level of aggregation is one that could usefully be investigated further.

The state-level results from the between estimator, in the bottom panel of Table 3, also show important differences from the county-level results. The predictive accuracy was lower, with R^2 values just below 0.70 with masked data products (or below 0.36 with unmasked data) compared to an R^2 of 0.86 at the county level. The elasticities were also lower at 0.84 compared to 1.05 in the county-level results with masked VNL data. Overall, this sensitivity to the level of spatial aggregation suggests a need to use findings from validation studies that are based on a similar level of spatial aggregation to what is used in ones' own study.

3.3. Results Using Earlier NTL Products

The V.2 VNL data products have only been recently available, so much of the literature has used older NTL data products such as V.1 VNL and DMSP stable lights composites. In this section, we examine how the results of estimating Equation (1) changed when older NTL data products are used. For comparisons, we used the V.2 VNL masked average radiance as that data product had the equal best performance in Table 2. Additionally, summing a (masked) mean to a county total is conceptually more consistent with GDP, which is the sum of economic activity in a county, than the case for summing a median.

In Table 4, we report estimates of Equation (1) for 2015–16 using either V.1 or V.2 VNL data as the right-hand side variable. For the analysis of temporal changes in GDP with respect to changes in NTL (the within estimator), V.2 is clearly superior, with an elasticity about four times larger (and an R^2 over 10-times larger). This is consistent with the expectation of the data creators, that the V.2 VNL series would do better at the analysis of lighting changes, due to using the same outlier removal threshold in all years rather than using a threshold that is year-specific, as in the V.1 VNL product [27]. Nevertheless,

we emphasize that the predictive power for county-level annual changes in GDP based on annual changes in NTL is very low, regardless of whether the V.1 or V.2 data are used. When cross-sectional differences were examined using the between estimator, performance of the V.1 and V.2 VNL data was very similar, with R^2 of about 0.86 and elasticities of about 1.03. Thus, existing cross-sectional results that have been established with the V.1 data should also hold with the V.2 data.

Table 4. Within and between estimators of GDP-lights elasticities: V.1 and V.2 VNL county-level results, 2015–16.

Indep Variables and Summary Stat	Within Estimator		Between Estimator	
	V.1 VNL	V.2 VNL	V.1 VNL	V.2 VNL
ln(sum of lights)	0.020	0.078 ***	1.037 **	1.026 ***
	(0.015)	(0.017)	(0.007)	(0.008)
R-squared	0.001	0.014	0.865	0.857

Notes: Based on a balanced panel of 3109 county-level units, observed in 2015 and 2016. The within estimator models include year and county fixed effects. The V.2 VNL product is the masked average radiance. Standard errors in parentheses (clustered at county level for the within-estimator), ** $p < 0.05$, *** $p < 0.01$.

Many studies of economic performance using NTL data continue to use DMSP data [16,33], even though the flaws in this data source, compared to VIIRS, have been known for almost a decade [23]. A key difference between these data sources is that even though the output grid for DMSP is only twice as coarse as for VNL (30 arc-seconds vs. 15 arc-seconds), the underlying spatial resolution of DMSP data is far coarser. This coarseness is due to geolocation errors [46], the smoothing of pixels into 5×5 blocks because onboard storage could not hold all the fine pixel data, and because there is no compensation for the expanded field-of-view as the Earth is viewed at an angle away from the nadir [24]. Consequently, the spatial precision of VNL images is at least 45 times greater than the precision of DMSP images [23]. One way that this imprecision shows up is through an exaggerated impression of urban extent from DMSP images [16,24,47].

Figure 1 shows how the lower 48 states of the U.S. (and also parts of Canada and Mexico) appear in the DMSP stable lights composite for 2013. Much of the land surface to the east of the 100° W meridian appears to be covered in light, and large clusters of light are also apparent around Denver, Salt Lake City, Phoenix, in California south of 39° N, and in Oregon and Washington north of 43° N. However, the picture shown with the V.2 VNL composite for 2014 appears very different, with cities having a far smaller lit area footprint than the DMSP data suggest (Figure 2). Notwithstanding the later overpass time of VIIRS, which may mean that some lights visible in the early evening have been turned off, the difference between Figures 1 and 2 reflects a key feature of DMSP of attributing city lights to places that are much less brightly lit (or even unlit). This feature contributes to noisy data that may distort apparent relationships between NTL and local economic activity.

There are several ways to numerically contrast Figures 1 and 2. A salient approach is to use spatial inequality statistics, as ever more studies use DMSP data to estimate inequality [39,40,48]. The overstated lit area in Figure 1 from DMSP blurring [24] makes it harder to distinguish areas of concentrated activity from other areas. Top-coding of DMSP data also attenuates differences between places. These spatially mean-reverting errors lead to far lower spatial inequality estimates when DMSP data are used, compared to when VIIRS data are used. When the Gini coefficient (an inequality measure that is zero for perfect equality and 1.0 for complete inequality) was calculated from the county-level GDP data, the average value over 2001–19 was 0.71 with no trend up or down. The V.2 VNL masked average radiances for 2014–19 gave a slightly lower value of 0.65, but it was not statistically significantly different to what the benchmark GDP data showed and also had no time trend. However, when the DMSP data for 2001–13 were used they gave an average Gini coefficient of just 0.50, significantly below the benchmark GDP estimate. Moreover, the DMSP data misleadingly suggested a downward trend in spatial inequality that was not apparent with the benchmark GDP data.
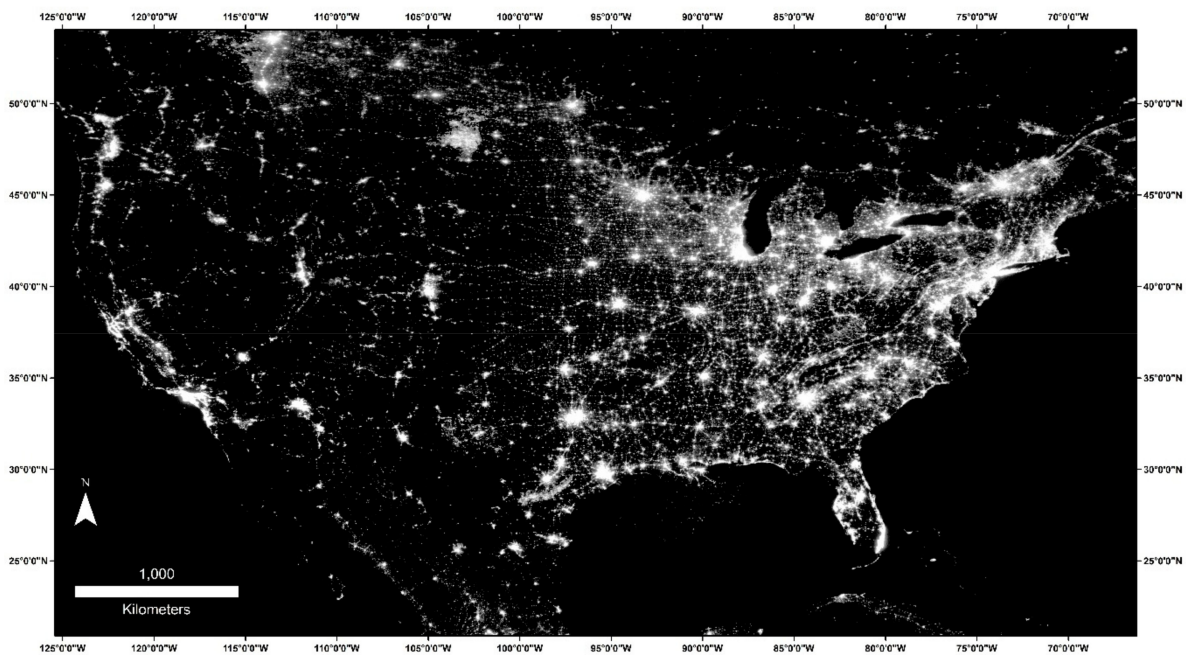
Figure 1. Night lights according to the DMSP stable lights annual composite, 2013.
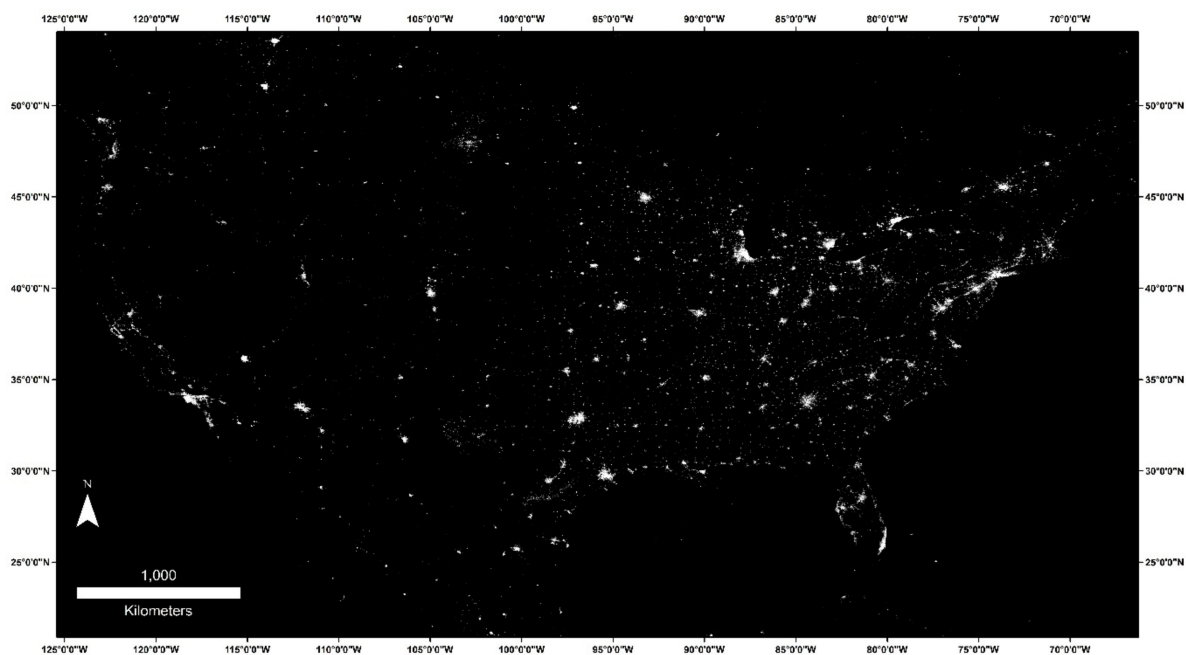

Figure 2. Night lights according to masked average radiance from the V.2 VNL, 2014.

In Table 5, we report the results of estimating Equation (1) using DMSP data for the panel of 3109 county-level units observed between 2001 (when the GDP data were first available) and 2013 (when the most widely used DMSP stable lights time-series ends). The table parallels Table 2, except for the earlier time period. For each year from 2001 to 2007, two DMSP satellites provided data (F14 and F15 through 2003, F15 and F16 through 2007). To deal with this extra information, we used three procedures reflecting approaches from applied studies. The first was to simply average the DN values from the two satellites operating in a particular year [49]; the second was to discard information from one satellite so that each year only had one source of data [13]; and the third recast the analysis in terms of satellite-years and introduced fixed effects for each satellite, in addition to fixed effects for each year [8]. The satellite-year approach creates an observation from the interaction of

a year and a satellite; for example, F15_2001 is a separate observation from F14_2001 or from F15_2002. Thus, when this method is used, the years with two satellites providing the data are counted twice as often as the years with just a single satellite. Therefore, to put equal weight on each year, the observations from 2001 to 2007 were weighted by 0.5 (as all of these years have two satellites providing the data) while a weight of 1.0 was used for the other years. Given that economics studies rarely use inter-calibrated DMSP data [50,51] as the year dummies in Equation (1) are claimed to deal with year-by-year fluctuations in the NTL time-series caused by sensor degradation and differences between satellites [7] we also did not use inter-calibrated DMSP data products.

Table 5. Relationships between DMSP NTL and county GDP: within- and between estimator results.

Independent Variables and Summary Statistics	Approach Used for Years with Two Satellites			Restricting to a 6-Year Time-Series (2008 to 2013)
	Averaging within Year	Use Observations of only 1 Satellite/Year	Use Satellite-Year Observations	
Within-estimator, for annual GDP changes within each county				
ln(sum of lights)	0.245 ***	0.173 ***	0.099 ***	0.190 ***
	(0.027)	(0.030)	(0.019)	(0.032)
Year fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
Satellite fixed effects	No	No	Yes	No
R-squared (Within)	0.100	0.070	0.042	0.080
Between-estimator, for average GDP differences between counties				
ln(sum of lights)	1.221 ***	1.222 ***	1.219 ***	1.208 ***
	(0.011)	(0.011)	(0.011)	(0.011)
R-squared (Between)	0.798	0.798	0.801	0.783
Sample size	40,408	40,408	62,163	18,653

Notes: Based on a balanced panel of 3109 county-level units, each year from 2001 to 2013. The within-year averaging affects years 2001 to 2007, which each have two satellites providing data. To use observations of only one satellite per year, we used F15 from 2001 to 2007, F16 in 2008 and 2009, and F18 from 2010 onwards. Standard errors in parentheses (clustered at county level for the within-estimator), *** $p < 0.01$.

How the issue of two DMSP satellites per year is dealt with affects the within-estimates of the GDP-lights elasticity, which can vary from 0.10 (using satellite-year observations) to 0.25 (using within-year averaging). A review of 18 economics studies using DMSP data found only two used satellite fixed effects while all used year fixed effects [16]. The results in Table 5 imply possible sensitivity of the results in this literature from not exploring other ways of incorporating multiple DMSP readings within a year (the within estimator is also affected by inclusion or exclusion of particular years, as seen below). This issue has no effect on the between estimator, which gives estimated elasticities of 1.22 across-the-board, because it is the same whether one first averages between satellites within a year and then averages over years, or instead averages over all satellite-years in one go.

Given the sensitivity to different ways of dealing with the observations from years with two DMSP satellites providing data, we also report the results in Table 5 for a 6-year time-series from 2008 to 2013. By necessity over this period, there is only one satellite available per year and so there is no sensitivity to different ways of dealing with multiple satellites in the same year. Additionally, these results (in the final column of Table 5) used a time-series that was of the same length as the time-series used for the V.2 VNL results shown in Table 2.

Two key patterns emerged from comparing the results in Table 5 with those in Table 2. First, the within estimator gave a higher GDP-lights elasticity using DMSP data for the period to 2013 than when using V.2 VNL data for the period since then, being about 50% higher if attention was restricted to the two 6-year time-series. Second, the between estimator showed that DMSP data gave elasticities more similar to those from the unmasked V.2 VNL data than those from the masked VNL data. Specifically, the estimated elasticity was 1.22 with DMSP data, 1.26 with unmasked V.2 VNL data, and only 1.05 with masked V.2

VNL data. In other words, the results with DMSP data were more like those coming from V.2 VNL data that had not had the background noise removed, which is an indirect way of saying that there is evidence of noise in the DMSP data. This noise reflects two features of DMSP data noted previously: attributing light to unlit places (blurring) and top-coding in brightly lit places [23,24]. Both features produce errors that cause a reversion toward the mean, and are likely to lead to elasticities being overstated rather than understated [35–37] if DMSP NTL data are on the right-hand side of regression equations.

The blurring and top-coding of DMSP that contribute to the noise in the NTL data are illustrated at finer scale in Figure 3, which maps four counties in western Massachusetts: Berkshire, Franklin, Hampshire, and Hampden using V.2 VNL data and DMSP data. The largest city in this region is Springfield (population: 160,000), and lights from this city (with masked average radiance exceeding 130 nW/cm^2/sr) are clearly visible in the middle of Hampden county in map (a) using V.2 VNL data for 2014. The largest cities in the other counties are far smaller, with populations of about 45,000 in Pittsfield (Berkshire Co.), 40,000 in Amherst (Hampshire Co.), and only 18,000 in Greenfield (Franklin Co.). The smaller size and lower brightness (e.g., no pixels in Pittsfield had an average radiance greater than 54 nW/cm^2/sr) of these other cities is also clear with the V.2 VNL data.

In contrast, the DMSP stable lights image for 2013 makes much of the area appear to be lit, with lights extending north from Springfield along Interstate 91 (I-91) corridor to Greenfield and into Vermont and New Hampshire (Figure 3b). Likewise, most of Berkshire county appears to be lit, with some parts seeming to be almost as brightly lit as Springfield. For example, Pittsfield has areas with DN = 60, which is almost as high as some areas in Springfield that have pixels with DN = 63, however, the reality seen in the V.2 VNL radiance data was that Pittsfield was only about 40% as brightly lit as Springfield, in line with being only one-quarter as populous.

When lights are aggregated to county level, the DMSP data greatly understate the differences between places. For example, the sum of lights for Franklin county was 35% of the sum of lights for Hampden county when DMSP data for 2013 were used. In contrast, the V.2 VNL data for 2014 showed that the sum of lights for Franklin county was just 9% of what was emitted by Hampden county. The GDP of Franklin county in either 2013 or 2014 was just 12% of that of Hampden county, and so the V.2 VNL data are a far more realistic proxy for what GDP reveals about the differences in economic activity in these two places.

This feature of DMSP data in understating differences between places is due both to blurring, which attributes light to unlit or less-lit places, and top-coding [33]. At least for the example of western Massachusetts, these two problems seemed to contribute equally to understated differences between places. In certain years (1996, 1999, 2000, 2002, 2004, 2005, and 2010), 'radiance-calibrated' DMSP data were derived from certain nights when NOAA asked the Air Force to turn down the amplification on the DMSP sensors, so that DN values were not top-coded in urban areas [52]. With these data for 2010, the sum of radiance-calibrated lights in Franklin Co. was one-quarter the sum of lights for Hampden Co., while the GDP of Franklin Co. in 2010 was only 13% of that of Hampden Co. In other words, the radiance-calibrated lights data made the smaller economy seem twice as large as what the GDP data showed. This improved over the three-fold overstatement of the smaller economy implied by the usual DMSP lights data, but the fact that the radiance calibrated lights still understated the GDP differences highlights the importance of the blurring problem in DMSP data, given that this problem is not dealt with by the radiance-calibration.

Features of DMSP data like blurring that contribute to exaggerated GDP-luminosity elasticities in between estimator results seem to hold in the extended DMSP series for the 2014–18 period. In Table 6, we report results using V.2 VNL data and extended DMSP data. The between estimator elasticity of 1.05 with V.2. VNL data was hardly changed from what was reported in Table 2 (as averaging was over five of the six years used in Table 2), but DMSP data for the same period gave an elasticity of 1.14. Once again, this exaggeration of the elasticity was consistent with mean-reverting errors in DMSP data. For the within

estimator results, the elasticity with DMSP data was smaller, perhaps because pre-dawn lights are less responsive to fluctuations in economic activity than are evening lights. For both the within and between estimators, the V.2 VNL data were more powerful predictors of GDP than were the DMPS data.
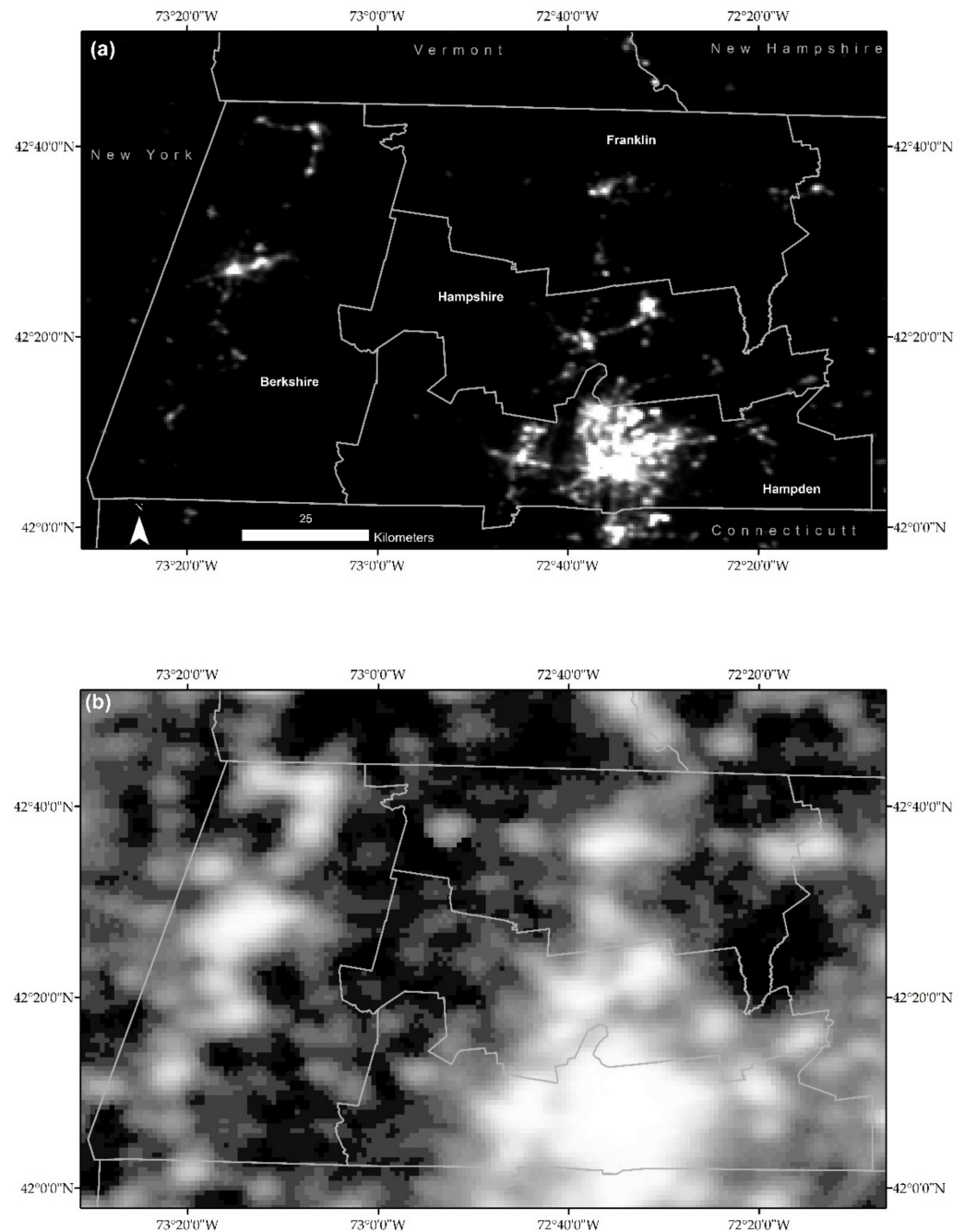

Figure 3. Night lights of western Massachusetts according to (**a**) V.2 VNL masked average radiance in 2014 and (**b**) DMSP stable lights in 2013.

A higher GDP-lights elasticity (for 2014–18) from V.2 VNL data than from extended DMSP data also holds with the country-level data. Recall from Table 1 (column 2) that the country-level elasticity with VNL data was 0.094 ± 0.038. This elasticity rose to 0.131 ± 0.034 when 2019 was omitted (so there is some sensitivity to sample periods). In contrast, with extended DMSP data, the elasticity was 0.063 ± 0.026 (the within R^2 was 0.046 compared to 0.118 with VNL data). Even noting that pre-dawn lights may vary less

with economic fluctuations than do evening lights, this is a far smaller GDP-lights elasticity than seen in prior results with DMSP data.

Table 6. Within and between estimators of GDP-lights elasticities: DMSP and V.2 VNL county-level results, 2014–18.

Indep Variables and Summary Stat	Within Estimator		Between Estimator	
	DMSP	V.2 VNL	DMSP	V.2 VNL
ln(sum of lights)	0.025 **	0.090 ***	1.139 ***	1.047 ***
	(0.010)	(0.018)	(0.011)	(0.008)
R-squared	0.004	0.016	0.767	0.862

Notes: Based on a balanced panel of 3109 county-level units, observed each year from 2014 to 2018. The within estimator models include year and county fixed effects. The V.2 VNL product was the masked average radiance. Standard errors in parentheses (clustered at county level for the within-estimator), ** $p < 0.05$, *** $p < 0.01$.

3.4. Results Using GDP by Industry

The U.S. has a larger share of GDP from the services sector than does any other major economy. The strength of the relationship between NTL and overall GDP depends on the structure of the economy because not all types of economic activity are equally reliant on lighting at night [25,26,41]. Thus, one way to examine how the above findings for the U.S. may apply to other countries is to look at estimates of Equation (1) that are disaggregated by industry, so that some extrapolation of the results to settings with different industrial structures can be considered.

The first two columns of Table 7 show that V.2 VNL data have higher predictive power for services sector economic activity than for goods-producing activities, whether examining cross-sectional differences or time-series changes. Hence, in countries where the services sector is less important than in the U.S., the NTL data may be less successful as a proxy for local GDP than they are in the U.S.

The private goods sector covers a range of industries and in some of them, there is a very weak, or entirely absent, relationship between NTL data and economic activity. The last two columns of Table 7 show the results for agriculture, forestry, fishing, and hunting (the primary sector), and for mining, quarrying and oil and gas extraction. The within estimator showed that changes in nighttime lights were not related to changes in primary sector economic activity, while they were only weakly related to changes in activity in the mining and oil and gas extraction sector. The between estimator results showed that GDP-lights elasticities were far smaller for these two industries than for all goods-producing industries and the R^2 values were much lower (and are almost zero for the primary sector).

Another way to consider the pattern shown in the third column of Table 7 is to divide counties into two groups, based on having an above-median or below-median share of agriculture in GDP (based on the 2014–19 averages). The within estimator results from column 3 of Table 2, where the elasticity was 0.12 ± 0.02, were re-estimated for these two sub-samples. In the counties where agriculture is more important, the elasticity was only 0.05 ± 0.02 (and the $R^2 = 0.01$), but where agriculture is less important, the elasticity was 0.18 ± 0.03 (and the $R^2 = 0.08$). Thus, NTL data may be less useful as a proxy for fluctuations in overall economic activity in places where agriculture is more important. Notwithstanding this result for fluctuations in economic activity, between estimator results in the first two columns of the lower panel of Table 8 suggest that V.2 VNL data remain a good proxy for differences in GDP between counties, whether they are more reliant on agriculture or not.

Table 7. Relationships between V.2 VNL masked average radiance and GDP by industry: counties, 2014–19.

Independent Variables and Summary Statistics	Services Sector	Private Goods Sector	Agriculture, Forestry, Fishing	Mining, Quarrying, Oil & Gas Extraction
	Within-estimator, for annual GDP changes within each county			
ln(sum of lights)	0.065 ***	0.154 ***	−0.038	0.161 ***
	(0.010)	(0.030)	(0.061)	(0.050)
Year fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
R-squared (Within)	0.013	0.008	0.000	0.002
	Between-estimator, for average GDP differences between counties			
ln(sum of lights)	1.097 ***	0.960 ***	0.136 ***	0.639 ***
	(0.010)	(0.010)	(0.020)	(0.032)
R-squared (Between)	0.813	0.747	0.016	0.130

Notes: Based on county-level panels, observed each year from 2014 to 2019, with N = 2935 cross-sectional units for the first two columns and N = 2850 cross-sectional units for the last two columns. The private goods-producing industries consist of agriculture, forestry, fishing, and hunting; mining, quarrying, and oil and gas extraction; construction; and manufacturing. Standard errors in parentheses (clustered at county level for the within-estimator), *** $p < 0.01$.

Table 8. Split-sample results for relationships between VIIRS V.2 VNL and county GDP.

Independent Variables and Summary Statistics	Agriculture Share of GDP		Population Density	
	Below Median	Above Median	Below Median	Above Median
	Within-estimator, for annual GDP changes within each county			
ln(sum of lights)	0.181 ***	0.053 ***	0.142 ***	0.093 ***
	(0.029)	(0.015)	(0.025)	(0.015)
Year fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
R-squared (Within)	0.080	0.005	0.020	0.040
	Between-estimator, for average GDP differences between counties			
ln(sum of lights)	1.073 ***	0.908 ***	0.799 ***	1.163 ***
	(0.011)	(0.012)	(0.012)	(0.012)
R-squared (Between)	0.849	0.800	0.726	0.851

Notes: Based on 3109 county-level units, observed each year from 2014 to 2019. The share of agriculture in GDP was averaged over all years and counties were then allocated into the above median or below median group based on the multi-year average. Population density was based on the 2010 census. Standard errors in parentheses (clustered at county level for the within-estimator results), *** $p < 0.01$.

One reason NTL data may be a less useful proxy for fluctuations in overall economic activity in more agricultural places is that there are some forms of non-agricultural activity, like retail shopping and wholesale distribution, which may occur at night aided by concentrated artificial light while this is less common for agriculture. Another factor is agriculture's use of space as a productive input, so population density and NTL intensity are lower in agricultural areas. For example, the counties with an above median share of agriculture in GDP had an average population density just under 40 people per square mile in the 2010 Census, while the counties with a below median share of agriculture had an average density more than 10-times higher, at almost 440 people per square mile.

The last two columns explore the role of population density more directly by splitting the sample into counties above and below the median density. In higher density counties, the predictive power of NTL data as a proxy for GDP was higher, for both the within estimator and the between estimator. The overall level and the composition of economic activity vary with population density, so relationships between NTL data and traditional indicators such as GDP will average over what could be quite disparate relationships for

particular places and types of activity, and this should be borne in mind when NTL data are used as a proxy.

4. Discussion

In this paper, we used a comprehensive and updated set of DMSP, V.1 VNL, and V.2 VNL nighttime lights data. We mainly examined the relationships with county-level and state-level economic activity for the U.S. over the 2001 to 2019 period, but we also provided some country-level results to link to the previous literature. Our motivation for using this rich set of NTL data products, and for using the lowest level spatial units that have GDP data available, stems from a concern that existing validation studies that assess NTL data as a proxy for economic activity are mainly for dated and imprecise DMSP data, and the most widely cited of these studies use aggregated spatial units such as nations or the first sub-national level. However, NTL data are increasingly used to proxy for economic activity at very local levels such as the third sub-national level and below. Another feature of recent applied studies is using NTL data to proxy for temporal fluctuations in local economies when evaluating the impacts of various shocks or policy interventions. In contrast, earlier studies tended to use NTL data to study regional differences in economic performance.

A key overall finding is that masked average radiance from the V.2 VNL data product was a better cross-sectional and time-series predictor of GDP than any of the other NTL products considered here (with the masked median also a good predictor). Masking to zero out background noise and ephemeral lights substantially improved predictive performance in cross-sections of county- and state-level GDP, and for time-series changes in county-level GDP. The masked V.2 VNL also better predicted time-series changes in GDP than did the V.1 VNL data, most likely because V.2 VNL uses a single multiyear threshold to isolate the background from lit grid cells while the year-by-year thresholds used for V.1 VNL may provide a less consistent basis for detecting changes. Comparisons with the predictive performance of extended DMSP data, which are based on pre-dawn readings from 2014 to 2018, also highlight the superiority of the masked V.2 VNL data.

When the various NTL data products faced the same benchmark GDP data, some predicted better than others. At least one reason for this is that some NTL data products are more error-ridden measures of true luminosity. The patterns of GDP-luminosity elasticities help to reveal the nature of these measurement errors. If either DMSP data or unmasked VNL data are used, the cross-sectional GDP-luminosity elasticity from the between estimator is exaggerated, with county-level estimates exceeding 1.20 (or 1.14 for the extended DMSP data product) compared with an elasticity of 1.05 from the masked VNL data that should have the least noise. This exaggeration of the elasticity suggests that measurement errors in DMSP data, and in unmasked VNL data, are mean-reverting rather than random. Consequently, these measurement errors will bias regression coefficients even if NTL data are the left-hand side variable, and can exaggerate coefficients rather than attenuate them if NTL data are the right-hand side variable.

There are at least two other consequences of mean-reverting errors in popular NTL data products like the DMSP annual composites. First, the literature that is beginning to use these data to estimate trends in spatial inequality may prove misleading, as inequality is significantly understated by DMSP data compared to what the GDP data and VIIRS data show. Second, attempts to splice together DMSP and VNL data to obtain a longer time-series face a key difficulty in finding an adjustment factor to make the DMSP data more like the VNL data. The measurement errors in DMSP data appear to vary with true but unknown luminosity; less brightly-lit areas have apparent luminosity overstated and more brightly-lit areas have it understated. Hence, no single adjustment factor, like an inter-calibration regression coefficient, can be most appropriate in all times and places. Moreover, spatial aggregation also affects the impacts of the measurement errors, as seen in the different patterns of results at county and state level.

The NTL data did far worse at predicting time-series changes in county GDP than at predicting in cross-sections of GDP. A prior study also found this in the V.1 VNL data [26],

but the results here are more compelling because they are from a longer time-series, using V.2 VNL data that should better measure lighting changes because they are derived from a constant threshold across years for isolating the background from lit grid cells. The weak relationship between changes in NTL and changes in GDP raises doubts about applied studies that show the effects of their treatment (e.g., a shock) on NTL data. If the GDP-luminosity elasticity is only 0.1 (and the within R^2 values are close to zero, as seen in Table 2), which is far lower than the elasticities in the literature reported from DMSP data at the national and first-subnational level, then it is hard to see how changes in NTL data are a good proxy for changes in local economic activity. In other words, estimates of the impact of the treatment on NTL data may not be very informative about the impact of the treatment on economic activity. In particular, treatment effects may be far smaller than presumed from econometric estimates using NTL data, especially if the researchers assume that cross-sectional elasticities hold in the time-series context [45].

5. Conclusions

There are several things that we can conclude from our analyses. First, masking to reduce measurement error improved the predictive power of V.2 VNL data. Second, predictive accuracy in county-level cross-sections was about 30-times higher than for county-level time-series changes in GDP. Third, the V.2 VNL data better predicted time-series changes in GDP than did the V.1 VNL data; likely due to V.2 VNL using a single multiyear threshold for isolating background from lit grid cells while the V.1 VNL uses year-by-year thresholds. Fourth, whether examined at the country level or county level, the relationship between recent temporal fluctuations in GDP and fluctuations in V.2 VNL data yielded a far smaller elasticity than was estimated when DMSP data were used for earlier years. Fifth, cross-sections of DMSP data provided similar results to what unmasked VNL data showed, indicating noise in the DMSP data (this pattern also holds if using the extended DMSP series). Relatedly, the DMSP data understate spatial inequality and the example we provide suggests that this comes in equal parts from blurring and top-coding.

The results reported here pertain to the United States—a setting where NTL data are not especially needed for research, given the abundance of other data on economic activity. However, the patterns of results across the various NTL data products for different spatial levels and for modeling time-series changes versus cross-sectional variation in economic performance should hold more broadly. For example, just using the U.S. data, it was possible to obtain a GDP-luminosity elasticity of 0.25 if a particular way of handling years with two DMSP satellites was used, which is quite close to the existing values in the literature beyond the U.S., despite more precise VNL data, suggesting an elasticity below 0.1. Moreover, the U.S. is a very diverse country, with types of economic activities in some places that are more like those in poorer countries. For example, given that NTL data are shown to be poor predictors of agricultural activity, or of changes in total economic activity in highly agricultural counties, there are grounds to question whether NTL data can be relied upon as a proxy for economic performance in predominantly agricultural settings in other countries. Relatedly, we also show that the NTL data were a less useful proxy for economic activity in less densely populated areas. Overall, our results suggest a need for greater caution in using NTL data as a proxy for economic activity, especially as findings from validation studies in different settings, or with different NTL data products, or at different levels of spatial aggregation may not translate to other settings.

Author Contributions: Conceptualization and methodology, J.G.; Software, validation, data curation and GIS analysis, G.B.-G.; Econometric analysis, preparation of original draft and revision editing, J.G.; Visualization, G.B.-G.; Project administration and funding acquisition, J.G. Both authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

References

1. Sutton, P.C.; Costanza, R. Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. *Ecol. Econ.* **2002**, *41*, 509–527. [CrossRef]
2. Ebener, S.; Murray, C.; Tandon, A.; Elvidge, C.C. From wealth to health: Modelling the distribution of income per capita at the sub-national level using night-time light imagery. *Int. J. Health Geogr.* **2005**, *4*, 1–17. [CrossRef]
3. Doll, C.N.; Muller, J.P.; Morley, J.G. Mapping regional economic activity from night-time light satellite imagery. *Ecol. Econ.* **2006**, *57*, 75–92. [CrossRef]
4. Elvidge, C.D.; Baugh, K.E.; Kihn, E.A.; Kroehl, H.W.; Davis, E.R.; Davis, C.W. Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *Int. J. Remote Sens.* **1997**, *18*, 1373–1379. [CrossRef]
5. Sutton, P.C.; Elvidge, C.D.; Ghosh, T. Estimation of Gross Domestic Product at sub-national scales using nighttime satellite imagery. *Int. J. Ecol. Econ. Stat.* **2007**, *8*, 5–21.
6. Henderson, V.; Storeygard, A.; Weil, D.N. A bright idea for measuring economic growth. *Am. Econ. Rev.* **2011**, *101*, 194–199. [CrossRef]
7. Henderson, J.V.; Storeygard, A.; Weil, D.N. Measuring economic growth from outer space. *Am. Econ. Rev.* **2012**, *102*, 994–1028. [CrossRef] [PubMed]
8. Chen, X.; Nordhaus, W.D. Using luminosity data as a proxy for economic statistics. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8589–8594. [CrossRef] [PubMed]
9. Nordhaus, W.; Chen, X. A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics. *J. Econ. Geogr.* **2015**, *15*, 217–246. [CrossRef]
10. Heger, M.P.; Neumayer, E. The impact of the Indian Ocean tsunami on Aceh's long-term economic growth. *J. Dev. Econ.* **2019**, *141*, 102365. [CrossRef]
11. Kocornik-Mina, A.; McDermott, T.K.; Michaels, G.; Rauch, F. Flooded cities. *Am. Econ. J. Appl. Econ.* **2020**, *12*, 35–66. [CrossRef]
12. Nguyen, C.N.; Noy, I. Measuring the impact of insurance on urban earthquake recovery using nightlights. *J. Econ. Geogr.* **2020**, *20*, 857–877. [CrossRef]
13. Eberhard-Ruiz, A.; Moradi, A. Regional market integration in East Africa: Local but no regional effects? *J. Dev. Econ.* **2019**, *140*, 255–268. [CrossRef]
14. Jagnani, M.; Khanna, G. The effects of elite public colleges on primary and secondary schooling markets in India. *J. Dev. Econ.* **2020**, *146*, 102512. [CrossRef]
15. Hodler, R.; Raschky, P.A. Regional favoritism. *Q. J. Econ.* **2014**, *129*, 995–1033. [CrossRef]
16. Gibson, J.; Olivia, S.; Boe-Gibson, G. Night lights in economics: Sources and uses 1. *J. Econ. Surv.* **2020**, *34*, 955–980. [CrossRef]
17. Li, G.; Cai, Z.; Liu, X.; Liu, J.; Su, S. A comparison of machine learning approaches for identifying high-poverty counties: Robust features of DMSP/OLS night-time light imagery. *Int. J. Remote Sens.* **2019**, *40*, 5716–5736. [CrossRef]
18. Gao, B.; Huang, Q.; He, C.; Ma, Q. Dynamics of urbanization levels in China from 1992 to 2012: Perspective from DMSP/OLS nighttime light data. *Remote Sens.* **2015**, *7*, 1721–1735. [CrossRef]
19. Rybnikova, N.A.; Portnov, B.A. Mapping geographical concentrations of economic activities in Europe using light at night (LAN) satellite data. *Int. J. Remote Sens.* **2014**, *35*, 7706–7725. [CrossRef]
20. Rybnikova, N.A.; Portnov, B.A. Estimating geographic concentrations of quaternary industries in Europe using Artificial Light-At-Night (ALAN) data. *Int. J. Digit. Earth* **2017**, *10*, 861–878. [CrossRef]
21. Lee, Y.S. International isolation and regional inequality: Evidence from sanctions on North Korea. *J. Urban. Econ.* **2018**, *103*, 34–51. [CrossRef]
22. Zhuo, L.; Ichinose, T.; Zheng, J.; Chen, J.; Shi, P.J.; Li, X. Modelling the population density of China at the pixel level based on DMSP/OLS non-radiance-calibrated night-time light images. *Int. J. Remote Sens.* **2009**, *30*, 1003–1018. [CrossRef]
23. Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proc. Asia-Pac. Adv. Netw.* **2013**, *35*, 62. [CrossRef]

24. Abrahams, A.; Oram, C.; Lozano-Gracia, N. Deblurring DMSP nighttime lights: A new method using Gaussian filters and frequencies of illumination. *Remote Sens. Environ.* **2018**, *210*, 242–258. [CrossRef]

25. Gibson, J.; Olivia, S.; Boe-Gibson, G.; Li, C. Which night lights data should we use in economics, and where? *J. Dev. Econ.* **2021**, *149*, 102602. [CrossRef]

26. Chen, X.; Nordhaus, W.D. VIIRS nighttime lights in the estimation of cross-sectional and time-series GDP. *Remote Sens.* **2019**, *11*, 1057. [CrossRef]

27. Elvidge, C.D.; Zhizhin, M.; Ghosh, T.; Hsu, F.-C.; Taneja, J. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sens.* **2021**, *13*, 922. [CrossRef]

28. Goldblatt, R.; Heilmann, K.; Vaizman, Y. Can Medium-Resolution Satellite Imagery Measure Economic Activity at Small Geographies? Evidence from Landsat in Vietnam. *World Bank Econ. Rev.* **2020**, *34*, 635–653. [CrossRef]

29. Elvidge, C.D.; Baugh, K.; Zhizhin, M.; Hsu, F.C.; Ghosh, T. VIIRS night-time lights. *Int. J. Remote Sens.* **2017**, *38*, 5860–5879. [CrossRef]

30. Baugh, K.; Elvidge, C.D.; Ghosh, T.; Ziskin, D. Development of a 2009 stable lights product using DMSP-OLS data. *Proc. Asia-Pac. Adv. Netw.* **2010**, *30*, 114. [CrossRef]

31. Jing, X.; Shao, X.; Cao, C.; Fu, X.; Yan, L. Comparison between the Suomi-NPP Day-Night Band and DMSP-OLS for correlating socio-economic variables at the provincial level in China. *Remote Sens.* **2016**, *8*, 17. [CrossRef]

32. Shi, K.; Yu, B.; Huang, Y.; Hu, Y.; Yin, B.; Chen, Z.; Chen, L.; Wu, J. Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: A comparison with DMSP-OLS data. *Remote Sens.* **2014**, *6*, 1705–1724. [CrossRef]

33. Gibson, J. Better night lights data, for longer. *Oxf. Bull. Econ. Stat.* **2021**, *83*, 770–791. [CrossRef]

34. Fuller, W.A. *Measurement Error Models*; Wiley: New York, NY, USA, 2009.

35. Bound, J.; Krueger, A.B. The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *J. Labor Econ.* **1991**, *9*, 1–24. [CrossRef]

36. Kim, B.; Solon, G. Implications of mean-reverting measurement error for longitudinal studies of wages and employment. *Rev. Econ. Stat.* **2005**, *87*, 193–196. [CrossRef]

37. Gibson, J.; Beegle, K.; De Weerdt, J.; Friedman, J. What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxf. Bull. Econ. Stat.* **2015**, *77*, 466–474. [CrossRef]

38. Bluhm, R.; Krause, M. *Top Lights: Bright Cities and Their Contribution to Economic Development*; Working Paper No. 2020-08; Monash University, SoDa Laboratories: Melbourne, Australia, 2020.

39. Lessmann, C.; Seidel, A. Regional inequality, convergence, and its determinants–A view from outer space. *Eur. Econ. Rev.* **2017**, *92*, 110–132. [CrossRef]

40. Singhal, A.; Sahu, S.; Chattopadhyay, S.; Mukherjee, A.; Bhanja, S.N. Using night time lights to find regional inequality in India and its relationship with economic development. *PLoS ONE* **2020**, *15*, e0241907. [CrossRef] [PubMed]

41. Keola, S.; Andersson, M.; Hall, O. Monitoring economic development from space: Using nighttime light and land cover data to measure economic growth. *World Dev.* **2015**, *66*, 322–334. [CrossRef]

42. Chen, X.; Nordhaus, W. A test of the new VIIRS lights data set: Population and economic output in Africa. *Remote Sens.* **2015**, *7*, 4937–4947. [CrossRef]

43. Henderson, J.V.; Squires, T.; Storeygard, A.; Weil, D. The global distribution of economic activity: Nature, history, and the role of trade. *Q. J. Econ.* **2018**, *133*, 357–406. [CrossRef] [PubMed]

44. Li, X.; Liu, S.; Jendryke, M.; Li, D.; Wu, C. Night-time light dynamics during the Iraqi civil war. *Remote Sens.* **2018**, *10*, 858. [CrossRef]

45. Asher, S.; Lunt, T.; Matsuura, R.; Novosad, P. Development research at high geographic resolution: An analysis of night-lights, firms and poverty in India suing the SHRUG open data platform. *World Bank Econ. Rev.* **2021**. [CrossRef]

46. Tuttle, B.T.; Anderson, S.; Elvidge, C.; Ghosh, T.; Baugh, K.; Sutton, P. Aladdin's magic lamp: Active target calibration of the DMSP OLS. *Remote Sens.* **2014**, *6*, 12708–12722. [CrossRef]

47. Small, C.; Pozzi, F.; Elvidge, C.D. Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sens. Environ.* **2005**, *96*, 277–291. [CrossRef]

48. Mirza, M.U.; Xu, C.; van Bavel, B.; van Nes, E.H.; Scheffer, M. Global inequality remotely sensed. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e1919913118. [CrossRef] [PubMed]

49. Storeygard, A. Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa. *Rev. Econ. Stud.* **2016**, *83*, 1263–1295. [CrossRef]

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

51. Hsu, F.-C.; Baugh, K.E.; Ghosh, T.; Zhizhin, M.; Elvidge, C.D. DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. *Remote Sens.* **2015**, *7*, 1855–1876. [CrossRef]

52. Elvidge, C.D.; Baugh, K.E.; Dietz, J.B.; Bland, T.; Sutton, P.C.; Kroehl, H.W. Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sens. Environ.* **1999**, *68*, 77–88. [CrossRef]