Noname manuscript No.
(will be inserted by the editor)

Can higher-quality nighttime lights predict sectoral GDP across subnational regions?

Urban and rural luminosity across the provinces in Türkiye

Author(s) Name

Received: date / Accepted: date

Abstract We investigate the potential of higher-quality nighttime light (NTL) data for predicting the sectoral gross domestic product (GDP) across subnational regions. Specifically, we use satellite images from the Visible Infrared Imaging Radiometer Suite (VIIRS) to study the relationship between regional luminosity and sectoral production across 81 provinces in Türkiye over the 2004-2020 period. Luminosity intensity is further decomposed into urban and rural components using land cover data. Employing pooled ordinary least squares, between-estimator, and within-estimator regressions, we examine the relationship between NTL and total GDP, agricultural GDP, non-agricultural GDP, industrial GDP, and services GDP. Our results show that urban NTL exhibits the most robust correlation with non-agricultural GDP. Notably, industrial GDP shows the highest GDP-NTL elasticity, pointing out the high predictive performance of urban NTL. From an estimation standpoint, the between-estimator yields superior predictive performance and more robust GDP-NTL correlations compared to other models. We also find that after accounting for spatial and temporal fixed effects, the within-estimator does not identify a significant relationship between NTL and GDP. Thus, we conclude by arguing that higher-quality nighttime light (NTL) data accurately predict sectoral GDP differences across regions but have weaker predictive power for annual GDP changes.

Keywords VIIRS-like data · sectoral GDP · nighttime light · land cover dave · Türkiye
JEL Classifications O15 · R12

Address(es) of author(s) should be given

1 Introduction

Nighttime light (NTL) detected by satellites serves as a valuable proxy for local development indicators in developing economies where GDP data are unavailable or limited. This proxy can be particularly useful when traditional economic statistics provided by the government are inconsistent in terms of definitions, time frames, and measurement instruments (Gibson and Boe-Gibson, 2021). Remotely sensed night light offers the advantage of objective measurements, regular updates, and broad coverage (Chen and Nordhaus, 2019). Originally, economists turned to NTL data derived from the Defense Meteorological Satellite Program's (DMSP) Operational Linescan System (OLS) as a means of proxying or predicting GDP (Chen and Nordhaus, 2011; Henderson et al., 2012; Lessmann and Seidel, 2017). However, this popular DMSP-OLS NTL data have flaws, including low spatial resolution (30 arc-second, i.e. approximately 1000 m), top-coding issues, and lack of calibration (Gibson et al., 2021). These limitations have threatened the accuracy of the analysis for areas with low population density and/or highly dependent on agricultural activities (Zhang and Gibson, 2022; Pagaduan, 2022). Despite its limitations, more than 150 economic studies have used DMSP-OLS NTL data, while other disciplines are swiftly transitioning to more advanced and accurate data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor (Gibson et al., 2020). Unlike blurred and saturated DMSP data, VIIRS data have higher resolution (15 arc-second, i.e. approximately 500 m) and do not suffer from over-saturation, blurring and lack of on-board calibration (Elvidge et al., 2017).

Nevertheless, NTL data, both from the DMSP and VIIRS, are not ideally suited for studying rural areas due to the type of economic activity prevalent in these regions. In rural sectors, economic activities such as agriculture emit less observable lighting, rendering them less detectable by NTL sensors. Conversely, NTL data are better at capturing economic activities in urban areas, including the construction, manufacturing, retailing, transportation, and service sectors (Chen and Nordhaus, 2019). This discrepancy is reflected in the urban-dominant nature of NTL observation, with few unlit pixels in urban sectors and almost no light emitted from croplands in rural areas during the night (Zhang and Gibson, 2022). Consequently, Keola et al. (2015) show that the relationship between NTL and GDP loses statistical significance in nations where agriculture contributes 50% or more of the total output. The authors also suggest utilising land cover data for predicting agricultural GDP, since vegetation mosaics and croplands exhibit a stronger association with the value added agriculture and forestry compared to NTL. Bundervoet et al. (2015) and Wang et al. (2019) explored alternative proxies like rainfall and vegetation to estimate economic activities in the agriculture sector. Furthermore, innovative approaches have emerged, such as integrating NTL datasets with day-time land cover data, as demonstrated by Pagaduan (2022).

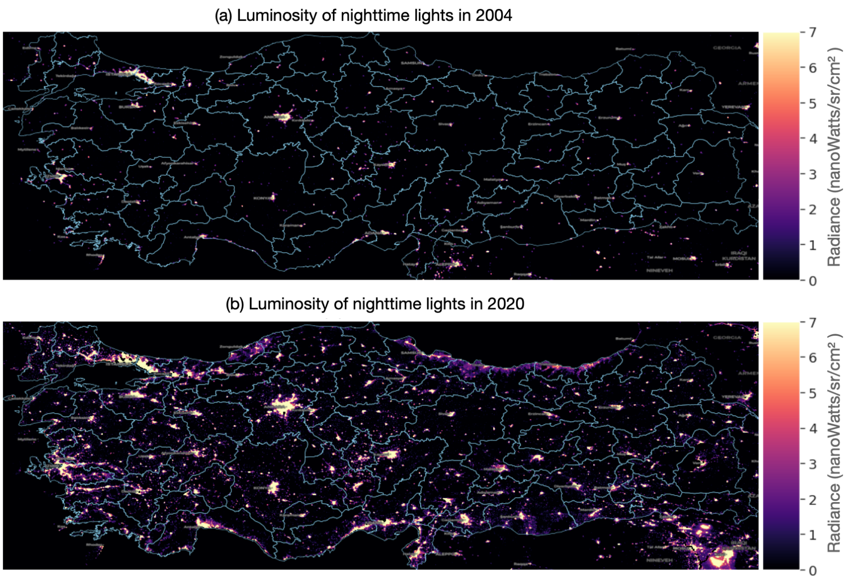

Fig. 1: Nighttime light images of Türkiye in 2004 and 2020

Notes: Luminosity intensity is measured in nanowatts per steradian per square centimeter (nanoWatts/sr/cm^2).

NTL images, such as those shown in Figure 1, have been used in a wide range of studies as a proxy for economic activity [1]. However, only a few studies have focused on estimating economic activities using NTL data for Türkiye. For example, using night lights, Basihos (2016) estimates provincial GDP and GDP per capita for the period 2001-2013 for Türkiye. Ustaoglu et al. (2021) estimates regional GDP in agricultural and non-agricultural sectors of Türkiye for the year 2015 by combining NTL data, land cover data, and GDP data. Hence, a key focus of this study is to examine whether the higher quality NTL data product can indeed serve as a reliable predictor of GDP in both urban and rural regions of Türkiye, during the period 2004–2020. Central to this investigation are inquiries into the distinct correlation between the newer VIIRS NTL intensity and GDP across various sectors. Addressing these questions constitutes the fundamental objective of our research.

This article contributes to the related NTL and economic literature in the following ways. First, we apply the novel NTL data developed by Chen et al. (2021) available from 2000 till now. This is an enhancement not only over the blurred and saturated DMSP data but also addresses the short temporal coverage of the NTL data from VIIRS processed by the Earth Observation Group (EOG), which is only available from April 2012 to the present. Previous studies aimed to create a temporally consistent and extended series by integrating the two data products: DMSP and VIIRS. However, these studies typically

[1] See Gibson et al. (2020) for a resent survey.

simulated DMSP-like products using existing monthly VIIRS composites, resulting in the inherited limitations of the DMSP data. In contrast, Chen et al. (2021) created a VIIRS-like product from existing DMSP annual composites and VIIRS monthly composites, beyond the traditional cross-sensor calibration approach. Second, following Pagaduan (2022), we decompose the NTL intensity into urban and rural areas by combining the NTL data with the MODIS daytime land cover dataset, since the DMSP or VIIRS NTL data poorly explain agricultural production in rural regions.Third, the literature on estimating economic activities using NTL data for Türkiye is rare. To our best knowledge, this paper is the first to estimate sectoral GDP using urban and rural NTL in Türkiye. NTL data can serve as a useful leading indicator of economic output, particularly in developing countries like Turkey, where income data at a regional scale are unavailable or limited. Thus, our paper shows that NTL can be a valuable indicator in terms of regionally disaggregated sectoral GDP.

The structure of this paper is organised as follows. Section 2 outlines the empirical methodology and remote sensing data used in this study. The results and discussion are presented in Section 3. Finally, Section 4 provides conclusions.

2 Methods and data

2.1 Panel data methods: Within and between variations

The primary goal of this study is to evaluate the predictive power of nighttime light (NTL) luminosity data as a proxy for official GDP data across various economic sectors. This evaluation will take into account both intra-regional and inter-regional differences in these two indicators. For this purpose, let us consider the following panel-data model:

$$\log(GDP)_{it} = \beta \log(NTL)_{it} + \mu_i + \varphi_t + \varepsilon_{it}, \tag{1}$$

where i denotes the regions, t denotes the years, μ_i is a region-specific effect, φ_t is a year-specific effect, and ε_{it} is a random error. Region-specific effects,μ_i, account for unobserved factors that remain constant over time. Year-specific effects, φ_t, account for unobserved factors that vary with time but are common across regions.

In this model, the key parameter is β, which summarizes the relationship between GDP and NTL. With the model defined in log terms, the value of β indicates the percentage increase in GDP corresponding to a one percent increase in NTL. The structure of Equation (1), however, does not immediately imply a cause-and-effect relationship. The role of the β parameter is solely predictive.

Various methods are available for estimating the parameter β. In this article, we focus on three common approaches:

$$\log(GDP)_{it} = \beta_{\text{Pooled}} \log(NTL)_{it} + \mu + \varepsilon_{it}, \tag{2}$$

$$\overline{\log(GDP)_i} = \beta_{\text{Between}} \overline{\log(NTL)_i} + \mu_i + \overline{\varepsilon_i}, \tag{3}$$

$$\log(GDP)_{it} - \overline{\log(GDP)_i} = \beta_{\text{Within}} \left[\log(NTL)_{it} - \overline{\log(NTL)_i}\right] + \varphi_t + \varepsilon_{it} - \overline{\varepsilon_i}, \tag{4}$$

A starting approach for calculating β relies on what is known as the "pooled' estimator, β_{Pooled}. When using this estimator (Equation (2)), we assume that the year-specific effects are null and all regions share a common intercept μ. This estimator also implies that the luminosity effect of a one percent difference between regions is equivalent to a one percent change within a single region. Consequently, its capability to assess the impacts of differences within and between provinces is constrained.

In contrast to the "pooled' estimator, the "between" and "within" estimators can be used to distinguish the NTL effects of differences between regions and over-time variations within those regions (Gibson and Boe-Gibson, 2021; Zhang and Gibson, 2022). Equation (3) presents the time-averaged (log) values of GDP and NTL. The coefficient β_{Between} reflects the impact on GDP when there is a variation in NTL between different regions. In Equation (4), we eliminate the unobservable region-specific factors (μ_i) from our estimates by deducting Equation 3 from Equation (1). The coefficient β_{Within} represents the impact on GDP resulting from variations in NTL within regions.

2.2 Study area: The 81 provinces of Türkiye

Türkiye stretches across the Anatolian peninsula in western Asia and Thrace in the Balkan region of southeastern Europe. According to the NUTS (Nomenclature of Territorial Units for Statistics) regional classification system, Türkiye has 81 provinces (NUTS-III), 26 sub-regions (NUTS-II), and 12 regions (NUTS-I). However, in Türkiye, like many developing countries, there is a huge subnational variation in levels of economic development across regions. The existence of significant and persistent disparities in economic development between Eastern and Western regions has been a central concern of economists and policymakers for decades.

Figure 2 shows the income distribution for the Turkish provinces in 2004 and 2020, which presents a similar pattern of NTL in Figure 1. We observe an east–west divide for both NTL and GDP, as expected. NTL intensity is higher in Western provinces compared to the less developed Eastern provinces, similar to GDP. The provinces with lower NTL intensity are generally clustered around the center and the east.

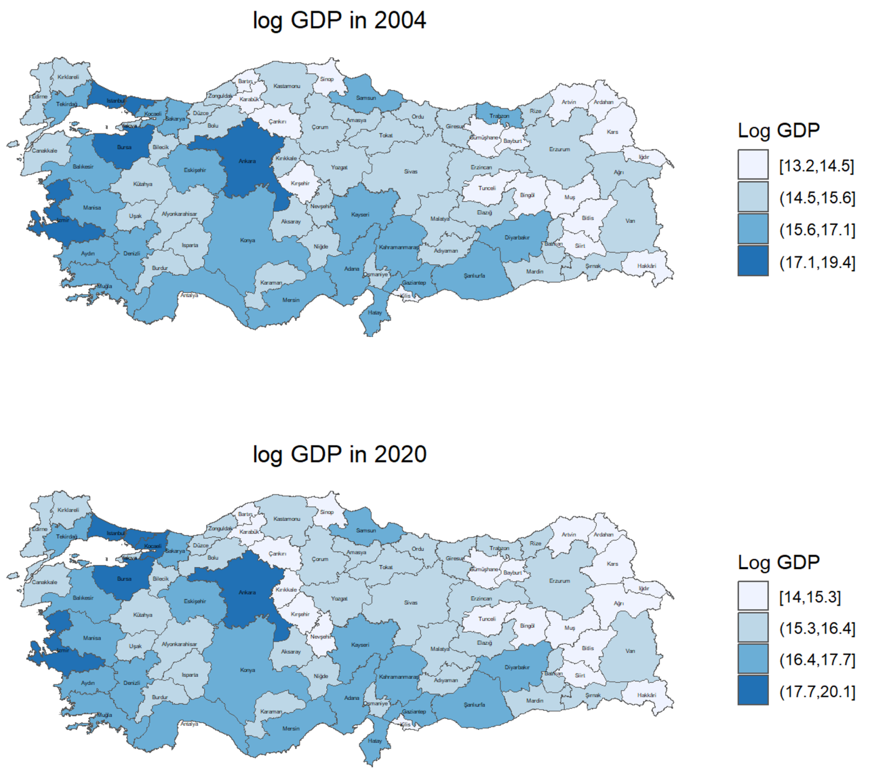

Fig. 2: Spatial distribution of GDP per capita in 2004 and 2020

Notes: In the maps, regions are classified into four categories based on the Fisher-Jenks optimization algorithm.

2.3 GDP and sectoral production

The dataset covers the 2004-2020 period for 81 provinces (NUTS-III level) in Türkiye. In this study, we use the recent official data of total regional GDP and sectoral regional GDP (in chain-linked volume index, (2009=100)). We obtain official GDP and sectoral GDP data from the Turkish Statistical Institute (TurkStat)[2]. GDP by provinces/regions in a chain-linked volume index is estimated by adjusting the inflation effect, enabling a more precise measurement of changes in output. In the empirical analysis, we use total regional GDP and sectoral GDP. We consider agricultural, non-agricultural, industry, and services GDP. Non-agricultural GDP comprises the combined value of industrial GDP and service sector GDP.

[2] https://www.tuik.gov.tr/

2.4 VIIRS-like nighttime lights

Nighttime light (NTL) satellite data, notably the DMSP-OLS NTL (2000-2012) and the VIIRS NTL composites (2013-2018), have emerged as valuable tools for investigating economic activities. However, the long-term analysis requires cross-sensor calibration of these two datasets due to differences in spatial resolutions and sensor design. To address this, an extended data set that spans 2000 to now, known as NPP-VIIRS-like NTL data, has been developed by Chen et al. (2021). The simulation process of this extended dataset featured the application of advanced image enhancement techniques, integrating a vegetation index and utilizing a convolutional neural network-powered autoencoder model. The cross-sensor calibration approach leverages deep learning technologies to transform calibrated DMSP data into VIIRS NTL series. Specifically, they utilise pixels from the calibrated DMSP data with digital numbers (DNs) equal to zero as a mask representing dark background areas in the simulated NTL series. Pixels in the simulated NTL series with intensities less than 1 nanowatt per square centimetre per steradian ($nWcm^{-2}sr^{-1}$) are adjusted to meet the detection limitation of the VIIRS sensor. Such post-processing procedures enable this extended dataset to exhibit excellent temporal consistency and spatial resolution similar to the VIIRS NTL product. This extended dataset is readily updatable and serves as a valuable proxy for tracking demographic and socioeconomic activities over a more extended time frame compared to the currently available products.

2.5 MODIS landcover classification

This study will use the MDC12Q1 product for land cover to divide urban and rural areas. The MODIS is a major sensor on the Terra and Aqua satellites of the National Aeronautics and Space Administration (NASA), capturing information from the entire earth's surface, such as vegetation, water bodies, and urban areas. The MDC12Q1 Version 6 product provides a combination of information on annual global land cover types after using supervised classification for reflectance data. Following Pagaduan (2022) and Keola et al. (2015), we will use land cover Class 12 (croplands), Class 10 (grasses/cereals) and Class 14 (croplands/natural vegetation mosaic) among the 17 International Geosphere-Biosphere Programme (IGBP) classifications for the agricultural sector. Class 13 (urban and built-up lands) refers to areas that account for at least 30% impervious surface area, including building materials, asphalt, and vehicles.

This study employs this land cover product to differentiate urban and rural areas, utilising the classifications to segregate NTL intensity with the urban-rural filter. This methodology aligns with the approaches of Pagaduan (2022) and Keola et al. (2015), focusing on specific classes for the agricultural and non-agricultural sectors. The geo-processing workflow is summarised in Fig. 3.

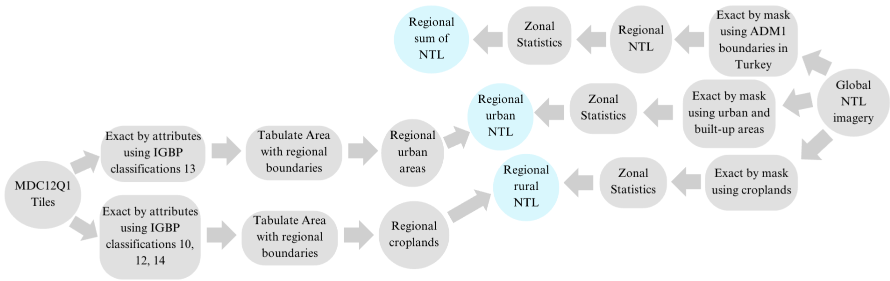

Fig. 3: Geo-processing workflow

Notes: This diagram is created by the authors, adapted from the workflow diagram in Pagaduan (2022).

3 Results and Discussion

Figure 4, 5 and 6 depict scatter plots of sectoral GDP and the sum of light in urban, rural, and both, using pooled OLS, between estimator, and within estimator [3]. Overall, NTL can be useful for predicting sectoral GDP, with a positive relationship with NTL. From the visualization of pooled OLS and between-estimator regressions in Figure 4 and Figure 5, non-agriculture GDP, including industry and service sectors, exhibits a stronger correlation with total or urban NTL compared to agriculture GDP and rural NTL. This observation is consistent with the conclusions drawn in most of the NTL literature. Urban economic activities, including industry, construction, manufacturing, and services, can be more accurately observed by nighttime light sensors (Keola et al., 2015; Chen and Nordhaus, 2019; Pagaduan, 2022). Either DMSP or VIIRS has limitations when capturing economic activities for areas with lower density specialising in agriculture (Gibson et al., 2021; Zhang and Gibson, 2022). However, the positive and significant relationship between NTL and GDP disappears after controlling the year and region fixed effects, as shown in Figure 6.

We put the regression results in Tables 1 and 2 [4] , using the same variables and regression methods as in Figures 4, 5 and 6. Leveraging the extensive temporal coverage of VIIRS-like NTL data, this study primarily concentrates on analysing the annual fluctuations within the time-series GDP. We present the results of the within-estimator regression based on time series variation in Table 1. After controlling the regional and year fixed effects, all the coefficients turn non-significant or negative, consistent with the visualisation in Figure 6. Therefore, all VIIRS-like NTL data in urban, rural and total areas poorly explain the annual changes in sectoral GDP in Türkiye. This result is similar

[3] Following the Frisch-Waugh-Lowell theorem, we visualise the two-way fixed effect regression using the residualized GDP and NTL.

[4] In the panel data analysis, we dropped three observations in certain regressions (e.g., the third column in Table 1), due to the 0 NTL values in the rural areas of Bartin and Rize in 2004 and 2005.

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
Sum of NTL (log)	-0.015			0.002			-0.019*			0.020		
	(0.012)			(0.028)			(0.011)			(0.024)		
Sum of NTL (log) in urban		-0.017			-0.007			-0.020**			0.001	
		(0.012)			(0.030)			(0.010)			(0.023)	
Sum of NTL (log) in rural			-0.005			0.008			-0.009			0.007
			(0.009)			(0.019)			(0.010)			(0.018)
Constant	15.113***	15.120***	15.034***	13.533***	13.598***	13.498***	14.839***	14.834***	14.747***	13.228***	13.377***	13.338***
	(0.093)	(0.084)	(0.060)	(0.221)	(0.213)	(0.118)	(0.088)	(0.071)	(0.065)	(0.187)	(0.163)	(0.116)
Observations	1,377	1,377	1,374	1,377	1,377	1,374	1,377	1,377	1,374	1,377	1,377	1,374
R-squared	0.020	0.012	0.034	0.052	0.038	0.056	0.010	0.004	0.025	0.081	0.038	0.048
Number of spatial units	81	81	81	81	81	81	81	81	81	81	81	81
Regional fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

Table 1: Relationship between provincial GDP and NTL (annual VIIRS-like) in urban, rural, and both: within-estimator result

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
	Pooled OLS											
Sum of NTL (log)	0.584***			0.662***			0.555***			0.332***		
	(0.014)			(0.017)			(0.013)			(0.011)		
Sum of NTL (log) in urban		0.582***			0.659***			0.554***			0.325***	
		(0.010)			(0.013)			(0.010)			(0.010)	
Sum of NTL (log) in rural			0.322***			0.371***			0.302***			0.229***
			(0.015)			(0.018)			(0.015)			(0.011)
Constant	10.119***	10.725***	12.806***	8.144***	8.840***	11.142***	10.011***	10.582***	12.593***	10.627***	11.024***	11.794***
	(0.124)	(0.083)	(0.127)	(0.153)	(0.105)	(0.150)	(0.120)	(0.080)	(0.122)	(0.104)	(0.078)	(0.088)
Observations	1,377	1,377	1,374	1,377	1,377	1,374	1,377	1,377	1,374	1,377	1,377	1,374
Number of spatial units	81	81	81	81	81	81	81	81	81	81	81	81
R-squared	0.572	0.707	0.240	0.533	0.657	0.231	0.565	0.700	0.231	0.383	0.456	0.250
	Between Estimator, for average GDP differences between provinces											
Sum of NTL (log)	1.040***			1.140***			1.003***			0.573***		
	(0.045)			(0.070)			(0.042)			(0.057)		
Sum of NTL (log) in urban		0.719***			0.798***			0.690***			0.394***	
		(0.040)			(0.054)			(0.038)			(0.043)	
Sum of NTL (log) in rural			0.528***			0.575***			0.507***			0.388***
			(0.092)			(0.110)			(0.088)			(0.061)
Constant	6.008***	9.629***	11.159***	3.831***	7.725***	9.509***	5.969***	9.489***	10.953***	8.455***	10.470***	10.520***
	(0.411)	(0.322)	(0.743)	(0.635)	(0.441)	(0.888)	(0.378)	(0.307)	(0.712)	(0.522)	(0.353)	(0.497)
Observations	81	81	81	81	81	81	81	81	81	81	81	81
R-squared	0.870	0.807	0.295	0.770	0.732	0.258	0.880	0.809	0.296	0.557	0.511	0.335

Note: Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: Relationship between provincial GDP and NTL (annual VIIRS-like) in urban, rural, and both: pooled OLS and between-estimator results

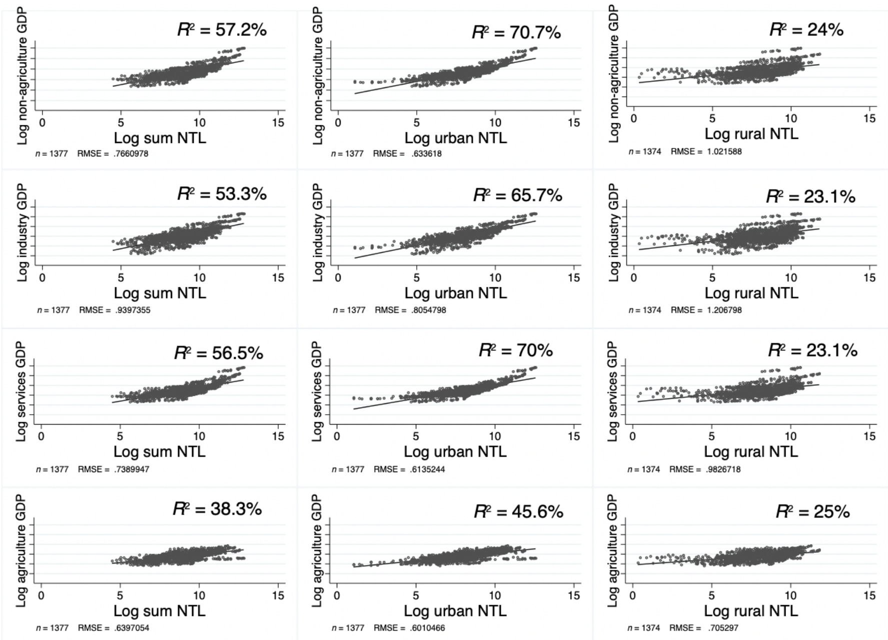

Fig. 4: Scatter plots of the pooled OLS regressions between NTL and sectoral GDP (2004-2020)

to what was found in VIIRS and DMSP in previous mainstream studies on NTL (Chen and Nordhaus, 2019; Gibson et al., 2021; Zhang and Gibson, 2022). This might imply that lights better provide a snapshot of economic activities at a specific point in time, rather than for short-term changes (Gibson et al., 2021), such as economic policies, natural disasters, and other external factors that can affect economic activities over time. The weak relationship between annual changes in NTL and annual changes in GDP also raises questions for applied research that explains the effects of treatment on changes in NTL, such as flooding or tsunami (Zhang and Gibson, 2022).

The results of between estimator is shown in Table 2, with the pooled OLS regression results as a reference. The R^2 for all dependent variables is two to three times higher when using total NTL or urban NTL as explanatory variables compared to rural NTL. Additionally, the R^2 values in columns (1) to (3) for non-agriculture sectors are approximately 1.5 times higher on average than those in column (4) for the agriculture sector. For instance, in columns (1) to (3) for non-agriculture sectors using the between-estimator regression, over 80% of the variation in non-agriculture and service production is explained by the variation in urban and total lights (with a R^2 above 80%).

In contrast, NTL variations provide a less effective explanation for variations in agricultural production, yielding an R^2 of 55.7% for total lights and 33.5% for rural lights. Therefore, total and urban NTL has stronger pre-

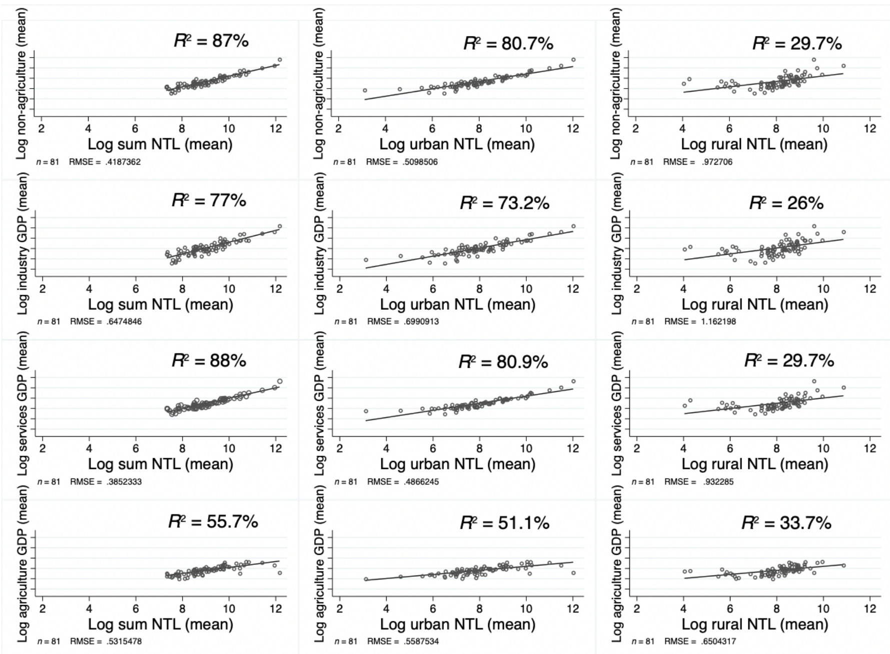

Fig. 5: Scatter plots of the between-estimator regressions between the cross-sectional NTL and time-averaged sectoral GDP (2004-2020)

dictability than rural NTL, especially for the non-agriculture sectors. Moreover, the R-squared values are larger when using the between estimator than the pooled regression, consistent with the findings reported in Gibson et al. (2021) for Indonesia. This suggests that VIIRS-like NTL is more useful in predicting cross-sectional time-averaged GDP between provinces, consistent with previous studies that utilised VIIRS and DMSP data (Keola et al., 2015; Chen and Nordhaus, 2019; Gibson et al., 2020; Gibson and Boe-Gibson, 2021).

All the coefficients in Table 2 are positive and statistically significant. The elasticity estimates in the between-estimator regression fall within the interval of values reported in studies analysing regional VIIRS NTL and GDP, such as Gibson and Boe-Gibson (2021) for the US, Gibson et al. (2021) for Indonesia, and Zhang and Gibson (2022) for China. For instance, in column (2) of the between-estimator regression, a 1% increase in the time-averaged total NTL corresponds to a 1.14% increase in the average production of the industry sector between provinces. Comparing all the coefficients of NTL from column (1) to (4), the cross-sectional time-averaged GDP in the industry sector has the largest elasticity of 1.14 with respect to the total time-averaged NTL data (column (2) in the between-estimator regression). This suggests that economic activities in the industry sector, including manufacturing and construction, are more easily detected through NTL (Pagaduan, 2022), despite lights explaining

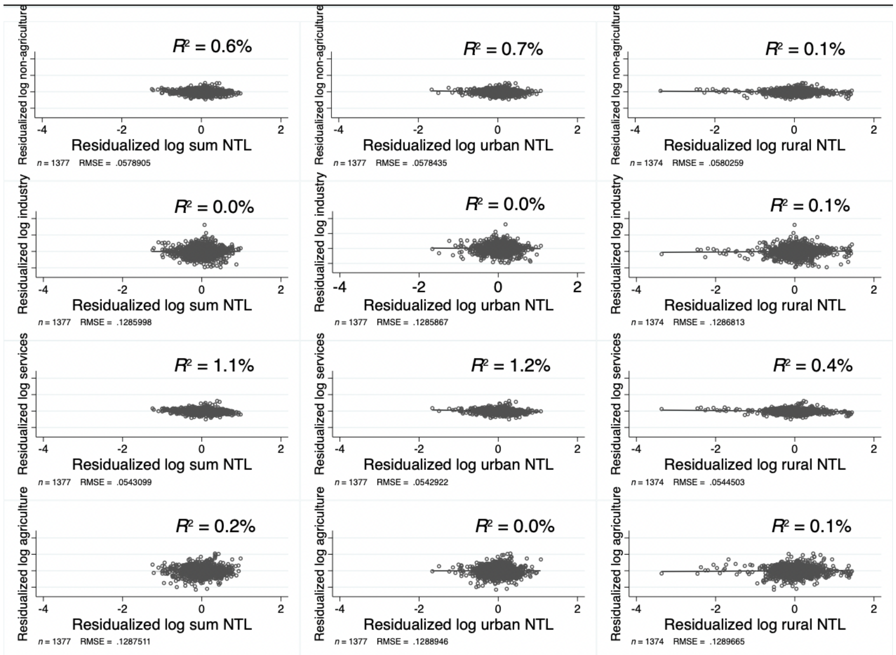

Fig. 6: Scatter plots of the residualised NTL and sectoral GDP in two-way fixed effect regressions (2004-2020)

a higher proportion of the variation in real GDP in both the non-agriculture and service sectors (columns (1) and (3)).

Beyond previous studies, our originality in Table 2 lies in separating the independent variable NTL into urban, rural and total areas for comparison. In the pooled OLS regression, the differences in predictive fitness between urban and total NTL align with Figure 4. R^2 values are higher when urban NTL is used as the explanatory variable in all sectors (columns 1 to 4 in the pooled OLS regression). There are no significant differences in the elasticities of GDP with respect to NTL between total and urban areas. However, the between-estimator results are slightly different. In all sectors, columns (1) to (4), the between-estimator elasticities of GDP with respect to total NTL are larger than those with respect to urban NTL. We also put the regression results using VIIRS and two versions of DMSP NTL as a robustness check, as shown in Appendix A and B.

4 Concluding remarks

In this paper, we examine the suitability of urban, rural and total nighttime light data for predicting differences in sectoral GDP between areas and for studying the temporal changes in sectoral GDP within subnational regions. We consider 81 provinces of Türkiye over the 2004-2020 period as a case study to evaluate the relationship between regional luminosity and sectoral production. Specifically, we evaluate higher-quality nitghttime lights (VIIRS-like satellite images) through three regression models: pooled OLS, between-estimator, and two-way fixed effect. Through the pooled regression, we observe better fitness of urban lights than the total lights with non-agriculture sectoral GDP. Both between-estimator and pooled regressions show that the predictability of NTL for non-agriculture sectoral GDP is approximately 1.5-fold that for the agriculture GDP. We found the largest GDP-(total) lights elasticity of 1.14 in the industry sector, which might imply that economic activities in the industry sector are more easily detected through NTL. Finally, the VIIRS-like NTL is a more powerful predictor for cross-sectional GDP than the time series GDP data in Türkiye, in line with the findings elsewhere. Therefore, the VIIRS-like data could be useful in predicting the time-averaged non-agriculture GDP for further regional studies in Türkiye. On the other hand, the weak predictability of NTL in the time-series GDP also poses challenges for applied research that interprets the effects of treatment on changes in NTL, as mentioned by Zhang and Gibson (2022).

References

Basihos, S. (2016). Nightlights as a Development Indicator: The Estimation of Gross Provincial Product (GPP) in Turkey.

Bluhm, R. and Krause, M. (2022). Top lights: Bright cities and their contribution to economic development. *Journal of Development Economics*, 157:102880.

Bundervoet, T., Maiyo, L., and Sanghi, A. (2015). Bright lights, big cities: measuring national and subnational economic growth in Africa from outer space, with an application to Kenya and Rwanda. *World Bank Policy Research Working Paper*, (7461).

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.

Chen, Z., Yu, B., Yang, C., Zhou, Y., Yao, S., Qian, X., Wang, C., Wu, B., and Wu, J. (2021). An extended time series (2000–2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration. *Earth System Science Data*, 13(3):889–906.

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.

Gibson, J. and Boe-Gibson, G. (2021). Nighttime Lights and County-Level Economic Activity in the United States : 2001 to 2019. (May).

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., and Boe-Gibson, G. (2020). Night lights in economics: Sources and uses. *Journal of Economic Surveys*, 34(5):955–980.

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.

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

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

Lessmann, C. and Seidel, A. (2017). Regional inequality, convergence, and its determinants–A view from outer space. *European Economic Review*, 92:110–132.

Pagaduan, J. A. (2022). Do higher-quality nighttime lights and net primary productivity predict subnational gdp in developing countries? Evidence from the Philippines. *Asian Economic Journal*, 36(3):288–317.

Ustaoglu, E., Bovkır, R., and Aydınoglu, A. C. (2021). Spatial distribution of GDP based on integrated NPS-VIIRS nighttime light and MODIS EVI data: A case study of Turkey. *Environment, Development and Sustainability*, 23(7):10309–10343.

Wang, X., Rafa, M., Moyer, J. D., Li, J., Scheer, J., and Sutton, P. (2019). Estimation and mapping of sub-national GDP in Uganda using NPP-VIIRS imagery. *Remote Sensing*, 11(2):163.

Zhang, X. and Gibson, J. (2022). Using multi-source nighttime lights data to proxy for county-level economic activity in China from 2012 to 2019. *Remote Sensing*, 14(5):1282.

Zhao, C., Cao, X., Chen, X., and Cui, X. (2022). A consistent and corrected nighttime light dataset (CCNL 1992–2013) from DMSP-OLS data. *Scientific Data*, 9(1):424.

Appendix

Appendix A: Robustness check using DMSP (2004-2013) and VIIRS NTL (2013-2020)

For the analysis during 2004-2013, we used the Consistent and Corrected Nighttime Lights (CCNL) dataset, a refined version of the DMSP OLS NTL. This dataset developed by Zhao et al. (2022) effectively mitigates inter-annual inconsistencies, data saturation, and blooming effects, thereby ensuring year-to-year comparability and improved data quality.

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
						Pooled OLS						
Log sum of NTL (DMSP)	1.199***			1.338***			1.152***			0.725***		
	(0.025)			(0.034)			(0.023)			(0.023)		
Log urban NTL (DMSP)		0.855***			0.954***			0.820***			0.440***	
		(0.013)			(0.020)			(0.012)			(0.017)	
Log rural NTL (DMSP)			0.551***			0.611***			0.527***			0.418***
			(0.032)			(0.039)			(0.030)			(0.022)
Constant	3.249***	8.474***	10.021***	0.554	6.380***	8.139***	3.366***	8.390***	9.891***	6.288***	10.055***	9.574***
	(0.247)	(0.105)	(0.302)	(0.338)	(0.156)	(0.366)	(0.233)	(0.100)	(0.288)	(0.232)	(0.135)	(0.206)
Observations	810	800	810	810	800	810	810	800	810	810	800	810
R-squared	0.746	0.840	0.271	0.660	0.750	0.237	0.753	0.844	0.271	0.548	0.461	0.314
					Between Estimator, for average GDP differences between provinces							
Log sum of NTL (DMSP)	1.356***			1.482***			1.313***			0.814***		
	(0.072)			(0.107)			(0.065)			(0.074)		
Log urban NTL (DMSP)		0.865***			0.960***			0.833***			0.444***	
		(0.041)			(0.061)			(0.039)			(0.053)	
Log rural NTL (DMSP)			0.587***			0.633***			0.568***			0.448***
			(0.107)			(0.130)			(0.102)			(0.072)
Constant	1.683**	8.389***	9.675***	-0.879	6.338***	7.933***	1.758***	8.292***	9.502***	5.396***	10.027***	9.291***
	(0.716)	(0.327)	(1.014)	(1.074)	(0.487)	(1.230)	(0.652)	(0.309)	(0.968)	(0.738)	(0.424)	(0.681)
Observations	81	80	81	81	80	81	81	80	81	81	80	81
R-squared	0.819	0.852	0.276	0.707	0.761	0.232	0.837	0.856	0.281	0.606	0.473	0.330

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

Table 3: Robustness check using DMSP NTL: pooled OLS and between-estimator results (2004-2013)

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
	Pooled OLS											
Log sum of NTL (VIIRS)	1.293***			1.386***			1.257***			0.715***		
	(0.020)			(0.031)			(0.017)			(0.024)		
Log urban NTL (VIIRS)		0.815***			0.894***			0.783***			0.425***	
		(0.020)			(0.020)			(0.014)			(0.017)	
Log rural NTL (VIIRS)			0.657***			0.683***			0.645***			0.529***
			(0.039)			(0.046)			(0.037)			(0.024)
Constant	0.847***	3.564***	3.981***	-0.103	2.726***	3.342***	0.842***	3.522***	3.859***	2.697***	4.301***	3.723***
	(0.091)	(0.057)	(0.165)	(0.141)	(0.078)	(0.197)	(0.080)	(0.054)	(0.157)	(0.108)	(0.066)	(0.104)
Observations	729	729	729	729	729	729	729	729	729	729	729	729
R-squared	0.854	0.817	0.285	0.739	0.741	0.233	0.878	0.820	0.300	0.560	0.476	0.396
Number of spatial units	81	81	81	81	81	81	81	81	81	81	81	81
	Between Estimator, for average GDP differences between provinces											
Log sum of NTL (VIIRS)	1.385***			1.488***			1.346***			0.760***		
	(0.053)			(0.088)			(0.044)			(0.072)		
Log urban NTL (VIIRS)		0.820***			0.901***			0.787***			0.426***	
		(0.043)			(0.059)			(0.041)			(0.050)	
Log rural NTL (VIIRS)			0.681***			0.711***			0.670***			0.552***
			(0.121)			(0.144)			(0.115)			(0.075)
Constant	0.428*	3.543***	3.876***	-0.572	2.698***	3.222***	0.436**	3.503***	3.755***	2.491***	4.297***	3.622***
	(0.243)	(0.172)	(0.520)	(0.407)	(0.235)	(0.620)	(0.204)	(0.164)	(0.494)	(0.331)	(0.199)	(0.322)
Observations	81	81	81	81	81	81	81	81	81	81	81	81
R-squared	0.897	0.821	0.285	0.782	0.748	0.235	0.921	0.823	0.300	0.587	0.481	0.407

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

Table 4: Robustness check using VIIRS NTL: pooled OLS and between-estimator results (2013-2020)

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
Log sum of NTL (DMSP)	0.001			0.057			-0.010			0.118***		
	(0.016)			(0.047)			(0.013)			(0.045)		
Log urban NTL (DMSP)		-0.029			0.032			-0.045***			0.054	
		(0.020)			(0.053)			(0.015)			(0.060)	
Log rural NTL (DMSP)			0.012			0.040			0.005			0.092***
			(0.011)			(0.032)			(0.008)			(0.033)
Constant	14.988***	15.242***	14.888***	13.000***	13.328***	13.186***	14.785***	15.053***	14.646***	12.242***	12.989***	12.543***
	(0.155)	(0.155)	(0.096)	(0.459)	(0.406)	(0.290)	(0.129)	(0.118)	(0.072)	(0.430)	(0.460)	(0.303)
Observations	810	800	810	810	800	810	810	800	810	810	800	810
R-squared	0.014	0.016	0.027	0.102	0.096	0.059	0.001	0.117	0.014	0.376	0.234	0.232
Number of spatial units	81	80	81	81	80	81	81	80	81	81	80	81
Regional fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

Table 5: Robustness check using DMSP NTL: within-estimator results (2004-2013)

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
Log sum of NTL (VIIRS)	0.051			0.059			0.052			0.189***		
	(0.059)			(0.105)			(0.059)			(0.066)		
Log urban NTL (VIIRS)		0.089			0.120			0.099			0.098	
		(0.068)			(0.111)			(0.063)			(0.075)	
Log rural NTL (VIIRS)			0.031			0.008			0.038			0.153***
			(0.042)			(0.078)			(0.040)			(0.054)
Constant	6.478***	6.365***	6.576***	5.928***	5.728***	6.154***	6.305***	6.154***	6.378***	5.094***	5.558***	5.302***
	(0.265)	(0.262)	(0.175)	(0.467)	(0.428)	(0.325)	(0.262)	(0.242)	(0.166)	(0.296)	(0.292)	(0.223)
Observations	729	729	729	729	729	729	729	729	729	729	729	729
R-squared	0.943	0.943	0.942	0.848	0.848	0.848	0.937	0.937	0.937	0.590	0.589	0.588
Number of spatial units	81	81	81	81	81	81	81	81	81	81	81	81
Regional fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

Table 6: Robustness check using VIIRS NTL: within-estimator results (2013-2020)

Appendix B: Robustness check using corrected DMSP data [5]

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
	Pooled OLS											
Log sum of NTL (corrected DMSP)	1.146***			1.274***			1.104***			0.678***		
	(0.025)			(0.034)			(0.024)			(0.024)		
Log urban NTL (corrected DMSP)		0.781***			0.871***			0.750***			0.388***	
		(0.012)			(0.018)			(0.011)			(0.016)	
Log rural NTL (corrected DMSP)			0.515***			0.569***			0.494***			0.399***
			(0.032)			(0.038)			(0.030)			(0.022)
Constant	3.255***	8.890***	10.112***	0.618*	6.854***	8.266***	3.342***	8.784***	9.968***	6.444***	10.382***	9.564***
	(0.264)	(0.098)	(0.317)	(0.358)	(0.146)	(0.383)	(0.247)	(0.092)	(0.302)	(0.246)	(0.130)	(0.215)
Observations	810	800	810	810	800	810	810	800	810	810	800	810
R-squared	0.719	0.844	0.245	0.631	0.751	0.213	0.730	0.849	0.246	0.507	0.432	0.296
	Between Estimator, for average GDP differences between provinces											
Log sum of NTL (corrected DMSP)	1.345***			1.459***			1.307***			0.788***		
	(0.072)			(0.109)			(0.064)			(0.076)		
Log urban NTL (corrected DMSP)		0.800***			0.883***			0.772***			0.395***	
		(0.037)			(0.056)			(0.034)			(0.050)	
Log rural NTL (corrected DMSP)			0.557***			0.596***			0.541***			0.434***
			(0.108)			(0.131)			(0.103)			(0.072)
Constant	1.185	8.735***	9.691***	-1.319	6.756***	7.999***	1.227*	8.609***	9.499***	5.297***	10.327***	9.215***
	(0.752)	(0.303)	(1.075)	(1.140)	(0.463)	(1.301)	(0.673)	(0.282)	(1.025)	(0.797)	(0.413)	(0.718)
Observations	81	80	81	81	80	81	81	80	81	81	80	81
R-squared	0.816	0.858	0.252	0.694	0.759	0.208	0.839	0.867	0.259	0.575	0.443	0.314

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

Table 7: Robustness check using the corrected DMSP NTL: pooled OLS and between-estimator results (2004-2013)

[5] The DMSP data after Pareto adjustment for top-coding by Bluhm and Krause (2022).

	(1) Non-agriculture GDP (log)			(2) Industry GDP (log)			(3) Service GDP (log)			(4) Agriculture GDP (log)		
Log sum of NTL (corrected DMSP)	0.010			0.046			0.002			0.091**		
	(0.016)			(0.045)			(0.013)			(0.038)		
Log urban NTL (corrected DMSP)		-0.021			-0.011			-0.025			-0.022	
		(0.024)			(0.072)			(0.017)			(0.068)	
Log rural NTL (corrected DMSP)			0.011			0.034			0.005			0.068**
			(0.010)			(0.027)			(0.008)			(0.034)
Constant	14.900***	15.187***	14.890***	13.086***	13.662***	13.225***	14.668***	14.904***	14.648***	12.456***	13.578***	12.724***
	(0.159)	(0.193)	(0.093)	(0.458)	(0.565)	(0.261)	(0.127)	(0.138)	(0.073)	(0.394)	(0.541)	(0.329)
Observations	810	800	810	810	800	810	810	800	810	810	800	810
R-squared	0.027	0.007	0.025	0.083	0.009	0.051	0.012	0.035	0.014	0.290	0.007	0.179
Number of spatial units	81	80	81	81	80	81	81	80	81	81	80	81
Regional fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

Table 8: Robustness check using the corrected DMSP NTL: within-estimator results (2004-2013)