

Regional growth, convergence, and spatial spillovers in India: A reproducible view from outer space

Carlos Mendez¹, Sujana Kabiraj², Jiaqi Li³

¹ Nagoya University, Nagoya, Japan

² Shiv Nadar University, Greater Noida, India

³ Nagoya University, Nagoya, Japan

Received: February 5, 2026/Accepted: September 19, 2026

Abstract. Using satellite nighttime light data as a proxy for economic activity, [Chanda, Kabiraj \(2020\)](#) studied regional growth and convergence across 520 districts in India. Adopting a reproducible open-science approach, this article builds on their work by extending their main findings on three empirical fronts. First, we illustrate regional convergence patterns using an interactive tool for satellite imagery visualization. Second, we assess the degree of spatial dependence in their main econometric specification. Third, we employ a spatial Durbin model to measure the role of spatial spillovers in the convergence process. Our results indicate that spatial spillovers increase the estimated speed of regional convergence. In our fully specified model, spillovers raise the convergence speed from about 3.0% to about 5.2% per year, which shortens the half-life of regional disparities from roughly 23 to 13 years. We close by illustrating how the same reproducible workflow extends beyond economic output, with an exploratory analysis of how luminosity relates to cultural participation across Indian states.

Key words: Regional convergence, Spatial dependence, Spatial Durbin model, Nighttime lights, Reproducible research, India

JEL classification: R11, R12, C21

1 Introduction

Regional growth and convergence matter most in federal states such as India. In such settings, spatial inequalities can threaten social cohesion and political stability. Consistent economic data, however, are rarely available below the state level. Satellite nighttime lights have relaxed that constraint by serving as a proxy for economic activity. Using those data, [Chanda, Kabiraj \(2020\)](#) document convergence across 520 Indian districts between 1996 and 2010. However, their analysis does not account for spatial spillovers. Such spillovers can arise from the diffusion of technology and from the movement of resources across neighboring districts.

This article extends that analysis along three dimensions. The first is replication: we adopt the data, the sample, and the conditioning variables of [Chanda, Kabiraj \(2020\)](#) without modification, so that any difference in the estimates reflects the spatial modeling alone. The second is substantive: we build an interactive visualization of regional

convergence, test formally for spatial dependence, and estimate a spatial Durbin model. The model shows that catch-up across Indian districts has a significant neighborhood dimension. The spatial spillover effect of initial luminosity is negative and statistically significant. Districts whose neighbors started from a lower base therefore grew faster than their own initial conditions alone would predict.

The third contribution is methodological, and it is the primary one. This article shows that a comprehensive spatial econometric study can be implemented using only open-source software and cloud computing. Our workflow covers every step, from visualizing the satellite imagery and constructing the neighborhood structure to testing for spatial dependence, modeling the spillovers, and running the robustness checks.

That workflow delivers the main empirical result of the article. Accounting for spatial spillovers raises the estimated speed of regional convergence from about 3.0% to about 5.2% per year. The implied half-life of regional disparities therefore falls from roughly 23 years to 13. The same workflow extends beyond economic output, and we illustrate that reach with an exploratory analysis of luminosity and cultural participation across Indian states. Our results suggest that luminosity is positively associated with media-based cultural consumption and negatively associated with community-based participation. This pattern is consistent with luminosity tracking electrification and urbanization rather than economic output alone.

Every figure and table in this article is produced by one of the six cloud-based computational scripts and notebooks listed in Table 1. The five analytical notebooks are written in Python and run in Google Colaboratory, so a reader needs no local installation and no licensed software: clicking *Run all* downloads the data and re-executes the analysis. The remaining notebook, N1, hosts the interactive visualization and its Google Earth Engine source code. The notebooks, the data, and the manuscript source are in the project repository at <https://github.com/quarcs-lab/project2025s-py>, and their rendered versions are embedded in the online edition at <https://quarcs-lab.github.io/project2025s-py/>.

Table 1: Computational notebooks and scripts underlying this article

Notebook	Content
N1. View from outer space	Interactive luminosity map with Earth Engine script
N2. Regional convergence	Luminosity growth on initial luminosity, with speed and half-life
N3. Spatial dependence	Weight matrices, Global Moran's I, and LISA cluster maps
N4. Spillover modeling	OLS and spatial Durbin estimates, impacts, and convergence speeds
N5. Robustness	Model 4 under six alternative weight matrices
N6. Spatial culture	Luminosity and cultural participation across Indian states

The rest of this article is organized as follows. Section 2 describes the data and the methods, and Section 3 presents the empirical results. Section 4 discusses the policy implications, the extension to multiple periods, the analysis of cultural participation, and new directions for research. Section 5 offers concluding remarks.

2 Data and methods

2.1 Data: Nighttime lights as a proxy for economic activity

Satellite images of nighttime light provide a proxy for economic activity below the national level. Henderson et al. (2012) pioneered this approach in economics. Their work exploits the empirical correlation between the intensity of artificial light observed from space and economic output on the ground. Chen, Nordhaus (2011) show that the proxy is most

informative where official statistics are of low quality. Nighttime light (hereafter NTL) data have therefore become a standard tool for the study of growth and convergence across national and subnational regions.

NTL data have been applied to a wide range of questions in development economics. One line of work studies the effect of decentralization on regional convergence ([Adhikari, Dhital 2021](#)). A second line adjudicates between national accounts and household surveys, and it finds that national accounts track the light data more closely ([Pinkovskiy, Sala-i-Martin 2016](#)). A third line estimates GDP per capita and spatial inequality at the global scale ([Lessmann, Seidel 2017](#)). The Indian literature is equally varied, and it covers welfare programs, colonial institutions, demonetization, and the pandemic.¹ [Chakravarty, Dehejia \(2017\)](#) document large regional disparities in India and caution that the goods and services tax may widen them further. Taken together, these studies show that NTL data are well suited to the analysis of subnational disparities.

Our empirical design follows [Chanda, Kabiraj \(2020\)](#). The dependent variable is the growth rate of nighttime lights per capita. The primary explanatory variable is the initial level of nighttime lights per capita. We construct both variables for 520 districts in India from data released by the National Geophysical Data Center (NGDC). These data record observations from the DMSP-OLS satellites over the period from 1996 to 2010. To mitigate top-coding in brightly lit areas, the NGDC also released a radiance-calibrated series for eight specific years, which combines high magnification settings for low-light regions with low magnification settings for bright ones. Six of those eight years fall inside our 1996–2010 sample window, and they are the dates our dataset holds. In this article, we use the radiance-calibrated series throughout.²

2.2 A reproducible workflow and interactive visualization

We implement the entire study in a reproducible workflow. Jupyter notebooks combine executable code, explanatory narrative, and computational output in a single document ([Kluyver et al. 2016](#)), which is the practical expression of literate programming ([Knuth 1984](#)). We document the data processing, the estimation, and the visualization in notebooks that run on a single Python kernel. Each result can therefore be traced to the code that produced it. Quarto then extends those notebooks into a manuscript framework ([Allaire et al. 2024](#)). Every figure and table is embedded from an identified notebook cell, and one source file generates the HTML, PDF, Word, and JATS XML versions without discrepancies between them. The toolchain is open source, and the data and code are public ([Peng 2011](#)). Each notebook also runs in the cloud on Google Colaboratory, which requires no local installation. Readers can therefore re-execute the whole pipeline in a browser and without proprietary software.

Satellite data have opened new dimensions of economic measurement ([Donaldson, Storeygard 2016](#)). Interactive visualization complements that measurement, because it reveals spatial and temporal heterogeneity that static maps obscure. Google Earth Engine lowers the computational barriers to such visualization ([Gorelick et al. 2017](#)). Its browser-based environment requires no local software, and its catalog already holds the DMSP-OLS and VIIRS nighttime lights collections ([Tamiminia et al. 2020](#)). Readers can therefore track light intensity across regions over time, and they can see where initially dim regions brightened fastest. That pattern is the visual counterpart of the convergence relationship estimated in Section 3.

¹Nighttime lights have been used in India to analyze the impact of public welfare programs ([Cook, Shah 2022](#)), the effects of colonial institutions ([Jha, Talathi 2024](#)), the impact of demonetization ([Chanda, Cook 2022](#)), and the effects of COVID-19 ([Beyer et al. 2023](#)).

²Following [Chanda, Kabiraj \(2020\)](#), our sample consists of 520 districts (out of a possible 593). There were 593 districts in 2001, which rose to 640 districts in the 2011 census. To match districts across the two census files, we merged newly split districts back with their parent districts in the 2011 census. Eight of the 47 new districts were created by splitting areas from multiple parent districts. Those new districts along with their multiple-origin districts were dropped from our sample. We dropped all the districts in the state of Assam, where more than half of the new districts were created in that manner. The 73 districts not in our sample are those of Assam and Tripura, the union-territory districts outside Delhi, the eight by which the nine districts of Delhi are collapsed to one, and 28 multiple-origin districts and their parents.

2.3 Regional convergence modeling

Neoclassical growth theory predicts that poor regions grow faster than rich ones. Diminishing returns to capital accumulation make the per capita growth rate of a region negatively correlated with its initial endowment (Solow 1956). Regions that share technology and preferences should therefore converge to a common steady state in the long run. We test that prediction within the growth regression framework of Barro, Sala-i-Martin (1992). Equation 1 states the convergence process in its simplest unconditional form.

$$\mathbf{g}_t = \beta_1 \mathbf{x}_{t-1} + \varepsilon_t \quad (1)$$

where $\mathbf{g}_t$ is an N -by-1 vector of per-capita NTL growth for the N regions over the period t . The vector $\mathbf{x}_{t-1}$ collects the initial (log) per-capita NTL of the same regions, and ε_t collects idiosyncratic error terms. The parameter β_1 measures the direction and the strength of regional convergence. A negative value implies that regions with lower initial NTL levels grow faster, which is the convergence hypothesis.

The estimated convergence coefficient translates into two quantities that are easier to interpret (Barro, Sala-i-Martin 1992). The first is the annual speed of convergence, and the second is the half-life. Because the dependent variable is average annual growth over T years, the implied convergence speed is $\lambda = -\ln(1 + \beta_1 T)/T$. The half-life is $\ln(2)/\lambda$, which is the time required to close half of the initial gap to the steady state. In our application $T = 14$ years. We apply the same transformation to the total effect of initial luminosity in the spatial models.

The unconditional framework contains no control variables, so it imposes a common steady state on every region. That assumption is restrictive, because regions differ in geography, in socioeconomic conditions, and in policy implementation. We therefore add state fixed effects, which absorb state-specific institutions and policies. We also add geo-climatic controls and district-specific conditions related to demographics, human capital, and infrastructure. Equation 2 summarizes the resulting conditional convergence framework.

$$\mathbf{g}_t = \beta_1 \mathbf{x}_{t-1} + \mathbf{X}_t \boldsymbol{\alpha} + \varepsilon_t \quad (2)$$

Here, the matrix $\mathbf{X}_t$ is an N -by- k collection of observations on control variables for each district in our sample, and it includes the state fixed effects. The vector $\boldsymbol{\alpha}$, with dimensions k -by-1, collects the associated regression coefficients. The conditioning set is adopted unchanged from Chanda, Kabiraj (2020). That choice is deliberate, because the present analysis extends that study rather than replacing it. Holding the controls fixed ensures that any difference between our estimates and their estimates comes from the spatial specification and not from a different choice of covariates.

The 16 district-level controls fall into three groups.³ Geography and agro-climate capture slow-moving determinants of the steady state that do not vary over the sample period. Demographics and human capital capture the factor endowments through which districts accumulate. Infrastructure captures access to the electricity grid and to markets. This last group matters here, because luminosity is itself partly a measure of electrification. The conditional models therefore compare districts that began the period with similar levels of grid access.

The demographic, human-capital, and electrification controls are measured in 1996, at the start of the growth period. They are therefore predetermined with respect to subsequent growth. The paved-roads variable is measured in 2000. The full list of control variables, with their measurement years and summary statistics, is given in Table A.1 of Chanda, Kabiraj (2020).

We estimate both specifications on a single long difference. The dependent variable is average annual growth between 1996 and 2010, and the regressors are measured at the

³Geography and agro-climate comprise agricultural suitability, rainfall, temperature, a malaria ecology index, terrain ruggedness, latitude, and distance to the coast. Demographics and human capital comprise the rural population share, log population density, the scheduled caste and scheduled tribe shares, the working population share, and the literate and higher-educated shares. Infrastructure comprises the share of households with an electricity connection and the log number of households with access to paved roads.

start of that period. This is the design of [Chanda, Kabiraj \(2020\)](#), and we retain it. The design identifies the average rate at which districts converged over fourteen years. It is silent, however, on how that rate evolved within the period. We return to this limitation, and to what a panel design would require, in the discussion.

2.4 Spatial dependence testing

Spatial dependence is the tendency for observations at nearby locations to be more similar than observations at distant locations. It can arise from spatial spillovers, from shared geographic conditions, or from inter-regional economic linkages ([Anselin 1988](#)). When it is present and left unmodeled, standard regressions yield biased parameter estimates and invalid inference. We therefore test for spatial dependence with the Global Moran's I statistic and with Local Indicators of Spatial Association (LISA). We apply both tests to the main variables of Equation 2.

The Global Moran's I statistic quantifies the overall degree of spatial clustering among geographic units ([Moran 1950](#)). It compares the deviation of each observation from the mean with the deviations of its neighbors. Equation 3 gives the formula.

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2} \quad (3)$$

where z_i is the deviation of observation i from the mean, and w_{ij} is the spatial weight that connects units i and j . The term n denotes the total number of observations. The statistic typically ranges from -1 to $+1$. Positive values indicate that similar values tend to be located near each other.⁴ We assess significance with a permutation-based inference approach.⁵

The global statistic does not reveal where the clusters and outliers are located. [Anselin \(1995\)](#) therefore proposed Local Indicators of Spatial Association, which decompose the global statistic into a contribution from each observation. Equation 4 defines the local statistic for observation i .

$$I_i = z_i \sum_j w_{ij} z_j \quad (4)$$

where the summation is over the neighbors of i , as defined by the spatial weight matrix. Each observation is then classified into one of four categories, according to how its value relates to the values of its neighbors. The High-High and Low-Low categories identify spatial clusters, and the High-Low and Low-High categories identify spatial outliers.⁶ We assess significance with conditional permutation tests. The LISA cluster maps report only the significant locations, typically at $p < 0.05$.

The spatial weight matrix $\mathbf{W}$ determines which districts are treated as neighbors, so its specification is central to the analysis. We employ a six-nearest-neighbor (6NN) matrix. It sets $w_{ij} = 1$ if district j is among the six nearest districts to i by centroid distance, and $w_{ij} = 0$ otherwise.⁷ Each district therefore has the same number of neighbors, even though district sizes vary widely across the country. We then row-normalize the matrix, as is standard practice, so that each row sums to one. The spatial lag $\mathbf{W}\mathbf{x}$ is then the average value among the neighbors of a district.

The choice of $k = 6$ balances two concerns. It is large enough to capture local spatial interactions, and it is small enough to exclude distant districts that are unlikely to exert economic influence. The data also support this choice. Among the seven weight matrices

⁴Values near zero suggest spatial randomness, and negative values indicate spatial dispersion, where dissimilar values are neighbors.

⁵The observed statistic is compared against a reference distribution generated by randomly reassigning values across locations.

⁶High-High (HH) indicates a high-value location surrounded by high-value neighbors, and Low-Low (LL) indicates a low-value location surrounded by low-value neighbors. High-Low (HL) is a spatial outlier where a high-value location is surrounded by low-value neighbors, and Low-High (LH) is the opposite spatial outlier.

⁷Centroid distances are measured in the geographic coordinates of the source file rather than in a projected metric, and rebuilding the matrix from a metric projection leaves 96.4% of the neighbor links unchanged.

considered in Table 4, the 6NN specification attains the lowest AIC, which indicates the best fit to the spatial structure of the data. We nevertheless treat the definition of the neighborhood as an open question. In Section 3 we re-estimate the preferred model under six alternative weight matrices, which range from contiguity-based schemes to inverse-distance schemes. Those results, reported in Figure 8 and Table 4, show that the direct and total convergence effects do not depend on this choice.

Evidence of significant spatial dependence would justify the use of spatial econometric methods. Ertur, Koch (2007) argue that the spatial Durbin model (SDM) accounts for such dependence in the convergence process. Fischer (2011) derives the same specification as the reduced form of a spatially interdependent growth model. The SDM also separates the direct effects of district characteristics from the indirect effects that operate through spatial channels (LeSage, Pace 2009).

2.5 Spatial spillover modeling

Our spatial spillover modeling builds on the spatially augmented Solow model of Ertur, Koch (2007), which was developed for countries. It also builds on the regional Mankiw-Romer-Weil extension of that model by Fischer (2011). Both models account for technological interdependence across economies. We apply this framework to an economy of N subnational regions. Each region is characterized by the Cobb-Douglas production function with constant returns to scale in Equation 5.

$$Y_{it} = A_{it} K_{it}^{\alpha_K} H_{it}^{\alpha_H} L_{it}^{1-\alpha_K-\alpha_H} \quad (5)$$

where Y_{it} is output, K_{it} is physical capital, H_{it} is human capital, L_{it} is the labor force, and A_{it} is the level of technological knowledge for region i at time t . The parameters α_K and α_H are the output elasticities with respect to physical and human capital. Technological knowledge is not exogenous in this framework, because it depends on internal and on external factors. Equation 6 states that relation.

$$A_{it} = \Omega_t k_{it}^\theta h_{it}^\phi \prod_{j \neq i}^N A_{jt}^{\rho W_{ij}} \quad (6)$$

The specification captures three components of technological progress:

- An exogenous component (Ω_t) that represents the common stock of knowledge across regions.
- An embodied component ($k_{it}^\theta h_{it}^\phi$) that reflects technology embedded in physical and human capital per worker.
- A spatial component ($\prod_{j \neq i}^N A_{jt}^{\rho W_{ij}}$) that captures technological interdependence between regions.

The third component is absent from the standard Solow model, and it is the source of the spatial spillovers in growth.

From this framework Ertur, Koch (2007) derived a spatial Durbin model that accounts for spillovers in the convergence process, with the human-capital terms following Fischer (2011). The model augments the conditional specification of Equation 2 with spatial lags of the dependent variable, of initial luminosity, and of the control variables, and can be written compactly as Equation 7:

$$\mathbf{g}_t = \beta_1 \mathbf{x}_{t-1} + \mathbf{X}_t \boldsymbol{\alpha} + \beta_2 \mathbf{W} \mathbf{x}_{t-1} + \mathbf{W} \mathbf{X}_t \boldsymbol{\gamma} + \rho \mathbf{W} \mathbf{g}_t + \boldsymbol{\varepsilon}_t \quad (7)$$

The vectors $\mathbf{g}_t$, $\mathbf{x}_{t-1}$ and $\boldsymbol{\varepsilon}_t$, the matrix $\mathbf{X}_t$, and the parameters β_1 and $\boldsymbol{\alpha}$ are defined as in Equation 2 above. The new terms are the spatial lags: $\mathbf{W} \mathbf{x}_{t-1}$ and $\mathbf{W} \mathbf{g}_t$ are N -by-1 vectors of neighborhood averages of initial luminosity and of growth, while $\mathbf{W} \mathbf{X}_t$ is the corresponding N -by- k matrix for the control variables. The coefficients β_2 , $\boldsymbol{\gamma}$ and ρ measure how strongly each of these neighborhood averages enters the growth of a district.⁸

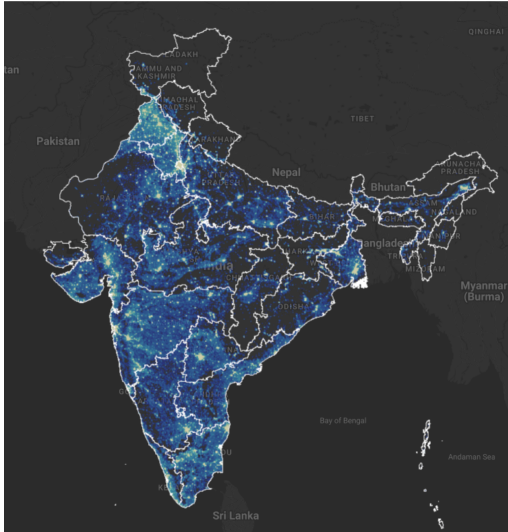
⁸We denote the spatial autoregressive parameter by ρ to avoid any confusion with λ , which denotes the implied speed of convergence throughout this article.

3 Results

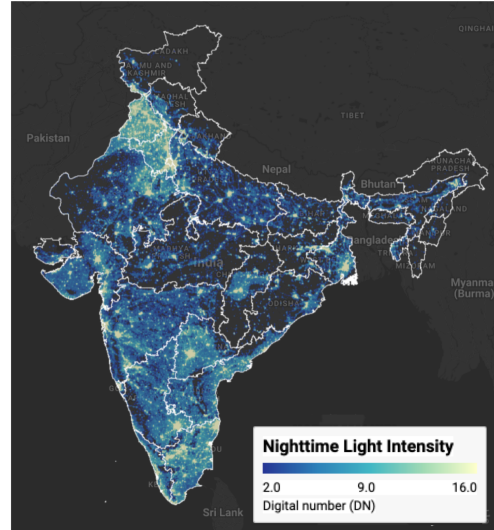
3.1 Regional convergence: An interactive exploration from outer space

Before presenting the regression results, we illustrate growth and convergence in nighttime lights visually, through an interactive map of India, a convergence scatterplot, and case studies of some of the poorest regions of the country.⁹

(a) Luminosity in 1996



(b) Luminosity in 2010



Source: Authors' visualization using pre-processed luminosity images from the Earth Observation Group (NOAA/NCEI). See [View from outer space](#) notebook for source code.

Notes: Luminosity is measured in radiance-calibrated digital number (DN) values from DMSP-OLS satellites. Interactive web application available at <https://bit.ly/india-rc-ntl>.

Figure 1: Regional luminosity in India: 1996 vs 2010

Figure 1 presents static maps of luminosity in the initial and final years, 1996 and 2010, captured from our interactive application. The maps show a noticeable increase in brightness across most of India by 2010, which, since nighttime lights proxy for economic activity, reflects economic growth over the period. Figure 2 then relates per-capita growth in nighttime lights to initial per-capita luminosity, and the relationship is inverse: the estimated β -convergence coefficient of about -0.02 implies an annual speed of convergence of roughly 2.3% and a half-life of about 30 years, consistent with the seminal finding of Barro, Sala-i-Martin (1992). Because β -convergence need not imply a narrowing of the distribution, we also check σ -convergence: the cross-district standard deviation of log per-capita luminosity falls from 1.09 in 1996 to 0.95 in 2010, a reduction of 13.5%, so spatial disparities did narrow over the period.

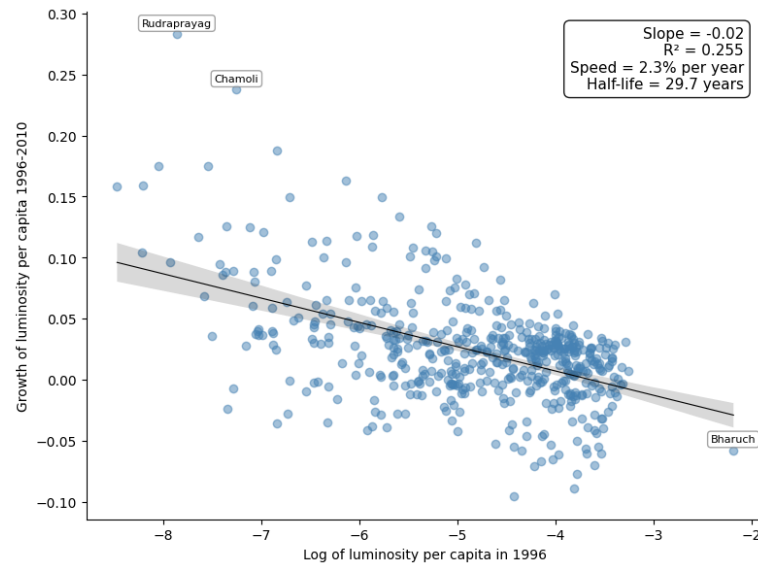
Figure 3 displays the change in luminosity over the study period in three economically disadvantaged states. Bihar is the poorest of them, with a per-capita income at 31.2% of the national average, while Uttar Pradesh and Chhattisgarh stand at 55.3% and 59.9% respectively.¹⁰ All three remain among the poorest states, but all three brightened noticeably over the period.

3.2 Spatial dependence is a feature of the convergence process

Before the formal econometric analysis, we examine the spatial distribution of the variables under study. Figure 4 maps initial luminosity in 1996 and the subsequent growth rate of luminosity over the 1996–2010 period across the 520 districts in our sample.

⁹The interactive web application is available at <https://bit.ly/india-rc-ntl>. It was developed using Google Earth Engine, and the source code is available in the [View from outer space](#) notebook.

¹⁰The figures are relative per capita income for 2000–01, the census year closest to the start of our study period, measured as per capita net state domestic product divided by per capita net national income. They are taken from Table 4 of Sanyal, Arora (2024), Working Paper EAC-PM/WP/31/2024.



Source: Data from Chanda, Kabiraj (2020). See Regional convergence notebook for source code.

Notes: Each point represents one of the 520 districts. The regression line shows the estimated beta-convergence relationship, and the annotation reports its slope, the R-squared, the implied annual speed of convergence, and the half-life. Outlier districts are labeled.

Figure 2: Regional luminosity convergence across districts in India

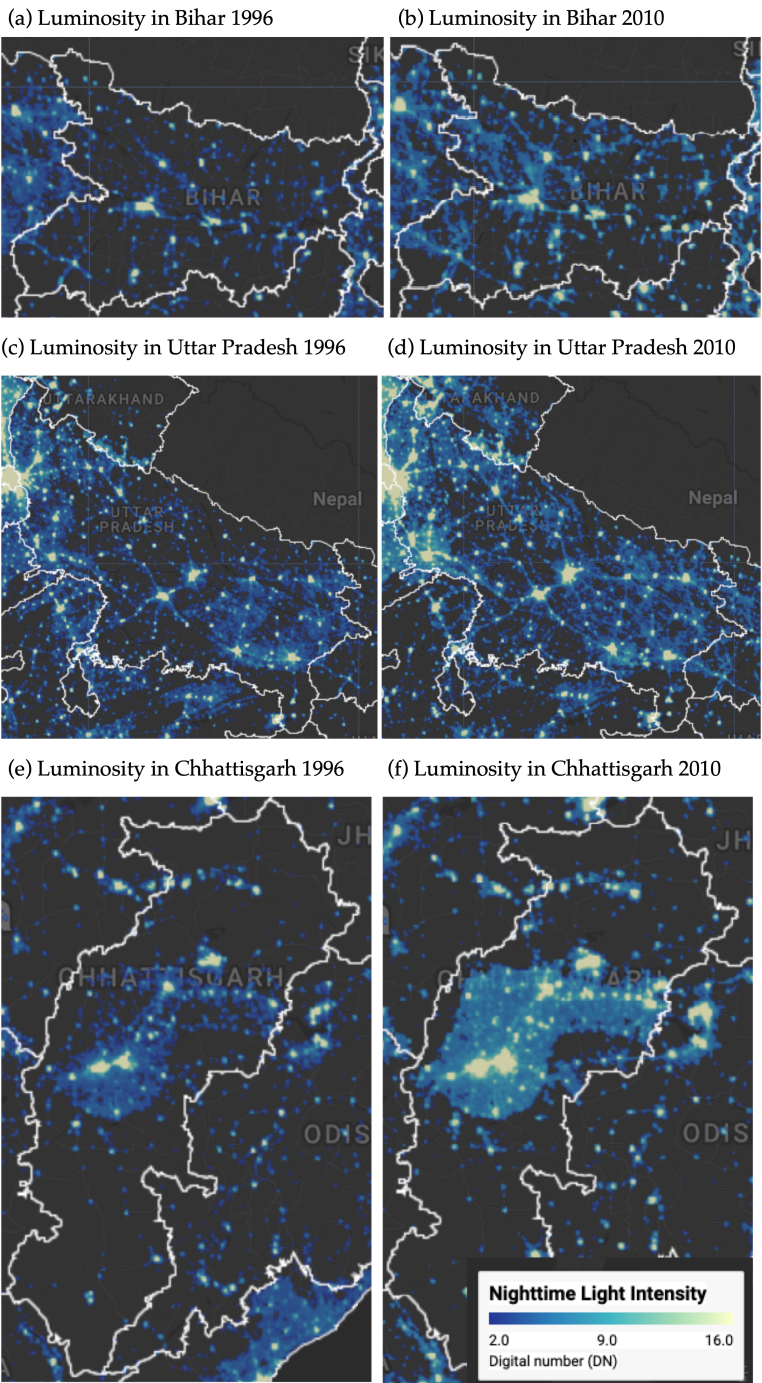
Panel (a) shows initial luminosity concentrated in specific geographic corridors, with higher values clustered in western and north-western India and markedly lower values across large parts of eastern and northern India. Panel (b) shows a pattern that is largely the inverse: districts that were initially bright grew more slowly, and districts that were initially dim grew faster. This spatial inversion provides a first visual indication that regional convergence in India may have an important spatial dimension.

A formal analysis of spatial dependence requires the neighborhood of each district, which the 6NN weight matrix described above supplies. Figure 5 visualizes that connectivity as a network, in which each node represents a district centroid and each edge connects a district to one of its six nearest neighbors.

Spatial dependence can be understood through two complementary concepts: the overall degree of clustering in the data, and the specific locations of statistically significant clusters and spatial outliers. The Global Moran's I statistic captures the first of these concepts. In our data, the Moran's I for initial luminosity per capita is 0.73 ($p = 0.001$), which indicates strong positive spatial autocorrelation: districts with high (low) initial luminosity tend to be surrounded by districts that are also bright (dim). The Moran's I for luminosity growth is 0.60 ($p = 0.001$), which confirms that the growth rates of neighboring districts are also significantly correlated.

We now turn to the LISA statistics, which locate the significant clusters and outliers behind these global measures. Figure 6 presents the Moran scatterplot and LISA cluster map for initial luminosity per capita. HH clusters mark the most luminous regions and their bright neighbors, while LL clusters highlight contiguous areas of low luminosity. These are predominantly in eastern and northern India, in Bihar, eastern Uttar Pradesh, and Jharkhand, together with the Himalayan north and the northeast.

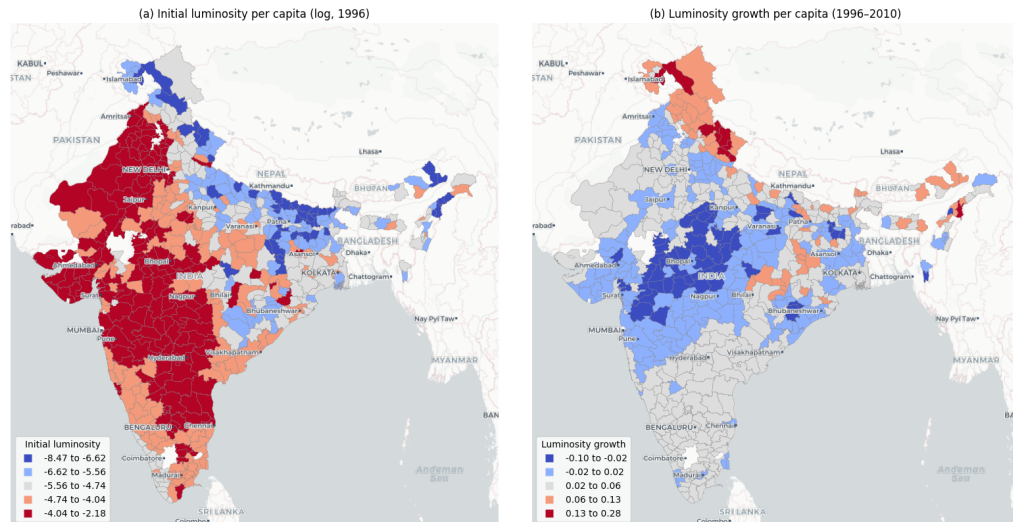
The same analysis applied to luminosity growth rates reveals a pattern that is largely the inverse of the initial clusters, as Figure 7 shows. No district that formed an HH cluster in initial luminosity forms one in luminosity growth: some appear instead as LL clusters, and most lose significance. Part of the initially dimmest cluster, in turn, emerges as an HH cluster in growth. That inversion is the spatial signature of the convergence process, and it motivates the use of spatial econometric methods.



Source: Authors' visualization using pre-processed luminosity images from the Earth Observation Group (NOAA/NCEI). See [View from outer space](#) notebook for source code.

Notes: Luminosity is measured in radiance-calibrated digital number (DN) values from DMSP-OLS satellites. Interactive web application available at <https://bit.ly/india-rc-ntl>.

Figure 3: Some illustrative examples of regional convergence



Source: Data from Chanda, Kabiraj (2020). See [Spatial dependence](#) notebook for source code.

Notes: Districts are classified into five categories using Fisher-Jenks natural breaks. Panel (a) shows log of luminosity per capita in 1996. Panel (b) shows luminosity growth per capita over 1996–2010.

Figure 4: Spatial distribution of initial luminosity and luminosity growth

3.3 Evidence of spatial spillovers in regional convergence

The regression results of Table 2 provide evidence of both unconditional and conditional convergence across districts in India, under ordinary least squares (OLS) and under the spatial specifications alike. The direct effects, which represent within-district convergence, remain negative and of similar order across specifications, ranging from -0.020 to -0.026. This consistency across specifications and estimation methods suggests that poorer districts are catching up to their wealthier counterparts, even after controlling for district characteristics and state-level fixed effects.

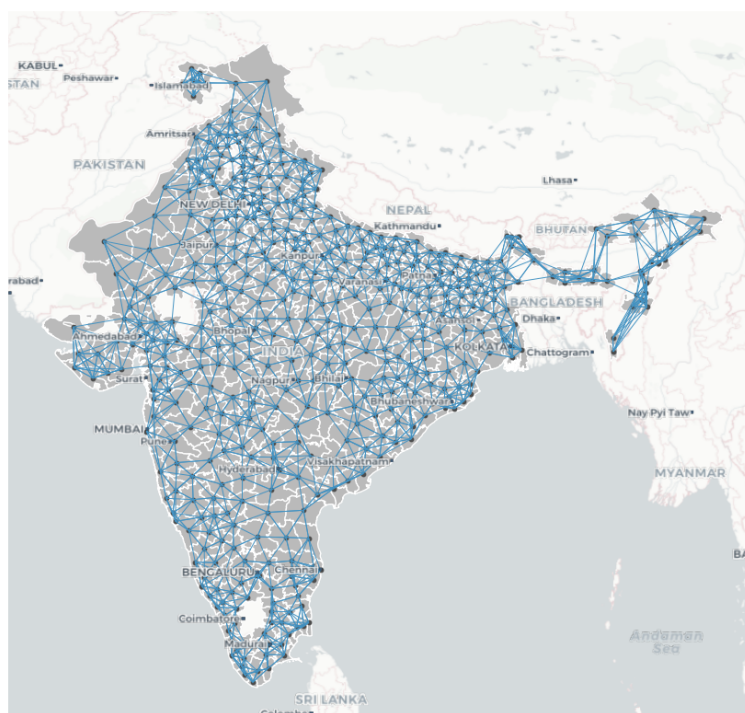
Table 2: Unconditional and conditional convergence across districts

	Model 1		Model 2		Model 3		Model 4	
	OLS	SDM	OLS	SDM	OLS	SDM	OLS	SDM
Direct	-0.020*** (0.002)	-0.021*** (0.002)	-0.022*** (0.003)	-0.021*** (0.002)	-0.025*** (0.003)	-0.026*** (0.002)	-0.025*** (0.003)	-0.025*** (0.002)
Indirect	—	-0.001 (0.006)	—	-0.001 (0.005)	—	-0.015* (0.008)	—	-0.012* (0.007)
Total	-0.020*** (0.002)	-0.022*** (0.006)	-0.022*** (0.003)	-0.022*** (0.005)	-0.025*** (0.003)	-0.041*** (0.009)	-0.025*** (0.003)	-0.037*** (0.007)
Controls	No	No	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	No	No	Yes	Yes
AIC	-1945	-2292	-2409	-2468	-2211	-2358	-2465	-2501

Notes: Heteroskedasticity-robust (HC1) standard errors for OLS and Monte-Carlo standard errors for the SDM impacts in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In the OLS specifications, the direct effect is -0.020 in Model 1 and -0.022 in Model 2, and it strengthens to -0.025 once district characteristics are added in Models 3 and 4. Omitting these characteristics therefore attenuates the estimated speed of convergence, and the spatial Durbin direct effects match this pattern in the conditional models, at -0.026 in Model 3 and -0.025 in Model 4. State fixed effects absorb unobserved state-level heterogeneity, such as differences in governance and institutional quality, which tightens the standard errors on the spatial impacts and moderates the spatial total effect rather than strengthening it further.

The spatial Durbin model also reveals spillover effects that OLS does not capture, and these indirect effects, which measure the influence of the initial conditions of neighboring districts on the growth rate of a district, follow a notable pattern across specifications. In



Source: Data from [Chanda, Kabiraj \(2020\)](#). See [Spatial dependence](#) notebook for source code.

Notes: Each node represents a district centroid. Each edge connects a district to one of its six geographically closest neighbors. The weight matrix is row-standardized.

Figure 5: Spatial connectivity structure based on six nearest neighbors

the unconditional models (Models 1 and 2), the estimated indirect effects are negative but small and statistically insignificant, at -0.001 in both.¹¹ Once district-level controls are introduced (Models 3 and 4), the indirect effects become both larger in magnitude and statistically significant at the 10% level, reaching -0.015 in Model 3 and -0.012 in our most comprehensive Model 4. The spatial channels of convergence therefore become discernible only after accounting for district-specific characteristics.

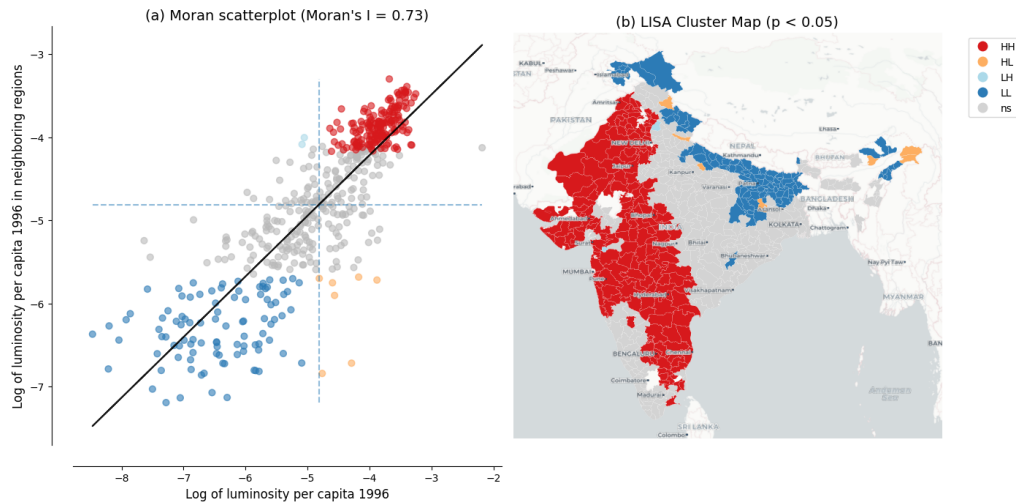
The total impact of initial conditions on growth, combining both direct and spillover effects, is substantially larger when we account for spatial dependence. In our fully specified model (Model 4), the total convergence effect in the spatial Durbin model (-0.037) is approximately 51% larger in magnitude than the OLS estimate (-0.025). Translated into an annual speed of convergence, this raises the implied speed from about 3.0% under OLS to about 5.2% under the SDM, and it shortens the implied half-life from roughly 23 to 13 years. The gap is even larger in Model 3 (SDM total -0.041 against OLS -0.025 , or 67%), but the Model 3 and Model 4 spillover estimates are statistically indistinguishable, so we anchor our interpretation on the preferred Model 4. These differences indicate that in this setting conventional non-spatial approaches understate the speed of regional convergence by failing to capture the additional convergence channels created through spatial spillovers.

Table 3 reports the implied speeds and half-lives for all four specifications. The spatial premium is concentrated in the conditional models. In Models 1 and 2 the OLS and SDM speeds are nearly identical, whereas in Models 3 and 4 the SDM raises the implied speed from 3.0% to 6.2% and 5.2%, shortening the half-life from 23 years to 11 and 13 years respectively.¹²

The Akaike Information Criterion (AIC) consistently favors the spatial Durbin model over OLS across all four specifications, and the most comprehensive model (Model 4

¹¹Without controls, the strong residual spatial dependence confounds the spillover estimates with omitted variables. The spatial autoregressive parameter also reaches about 0.80 in Model 1, so its decomposition is the most sensitive of the four to how the spatial multiplier is computed.

¹²The unconditional speeds are 2.3% against 2.6% in Model 1 and 2.6% against 2.7% in Model 2.



Source: Data from Chanda, Kabiraj (2020). See [Spatial dependence](#) notebook for source code.

Notes: Panel (a) shows the Moran scatterplot with Global Moran's I statistic. Panel (b) shows the LISA cluster map with statistically significant clusters at $p < 0.05$ based on 999 permutations.

Figure 6: Spatial dependence in the initial level of luminosity

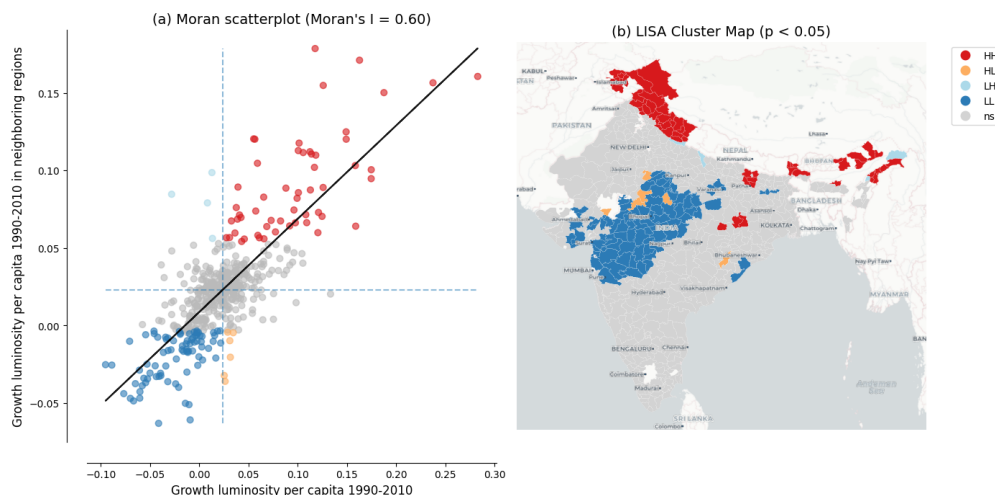
Table 3: Implied annual speed of convergence and half-life by model, from the OLS and SDM total effects of initial luminosity

Model	OLS speed	SDM speed	OLS half-life	SDM half-life
Model 1	2.3%	2.6%	30 yr	27 yr
Model 2	2.6%	2.7%	27 yr	26 yr
Model 3	3.0%	6.2%	23 yr	11 yr
Model 4	3.0%	5.2%	23 yr	13 yr

SDM) achieves the lowest AIC value of -2501 compared to -2465 for the corresponding OLS.¹³ These results suggest that regional convergence in India operates not only through district-specific factors but also through spatial interactions between neighboring districts, so that ignoring those interactions yields both underestimated convergence speeds and inferior model performance.

A word of caution is in order about how the indirect effects should be read. The spatial Durbin model identifies spatial dependence in reduced form: it establishes that the growth of a district covaries with the initial conditions of its neighbors, conditional on its own characteristics, but it does not establish why. Several distinct processes generate the same reduced-form pattern, since neighboring districts share geographic and agro-climatic conditions, are often governed by the same state and so inherit common institutional and policy environments that our state fixed effects absorb only in part, are connected by infrastructure whose effects do not stop at administrative boundaries, and are exposed to any omitted determinant of growth that is itself spatially clustered. The measurement issues discussed below point in the same direction, since blooming spreads light across district boundaries and can raise the apparent correlation between neighbors for reasons unrelated to economic interaction. We therefore read the indirect effects as evidence that spatial structure is a feature of the convergence process, and not as an estimate of the causal effect of the conditions of one district on the growth of another. Establishing which channels are at work, and whether they are causal, would require research designs built for that purpose, a point we return to in the discussion.

¹³The improvement from incorporating spatial structure is most pronounced in the unconditional specification, where the AIC drops by 347 units when moving from OLS to SDM in Model 1, reflecting the large amount of spatial dependence left unmodeled by ordinary regressions.



Source: Data from Chanda, Kabiraj (2020). See [Spatial dependence](#) notebook for source code.

Notes: Panel (a) shows the Moran scatterplot with Global Moran's I statistic. Panel (b) shows the LISA cluster map with statistically significant clusters at $p < 0.05$ based on 999 permutations.

Figure 7: Spatial dependence in the growth rate of luminosity

3.4 How these estimates compare with the convergence literature

The speeds reported above can be placed against three benchmarks. The first is other work on Indian districts using the same proxy. Beyer, Yamada (2020) report 2.4% absolute and 3.7% conditional convergence for India over 2013–2019, against our own 2.3% and 3.0%. That two independent applications of the same measurement strategy recover similar non-spatial rates, despite different periods and different generations of the sensor, is reassuring about the proxy.

The second benchmark is convergence measured from luminosity elsewhere. Basher et al. (2023) find absolute convergence of 4.57% per year across the 64 districts of Bangladesh, roughly double the 2.3% we obtain on the same unconditional basis. Carington, Jiménez-Ayora (2021) report a speed near 2% for Ecuadorian provinces, and Kacou (2022) finds multiple equilibria across 49 African countries. Our own unconditional non-spatial estimate therefore sits near the familiar 2% regularity, while the conditional spatial estimates lie well above it.¹⁴ A meta-analysis of some 600 published estimates nevertheless warns against treating that regularity as a law (Abreu et al. 2005). These comparisons locate our estimates rather than validate them.

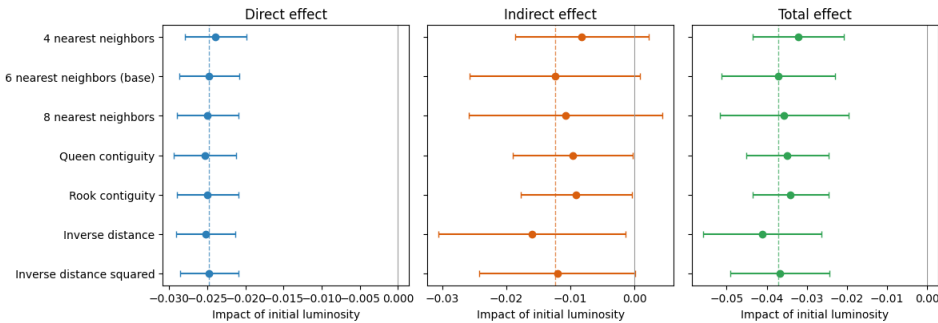
The third benchmark bears most directly on what this article adds. That convergence estimates are sensitive to spatial dependence is established, but the direction of the adjustment is not. Our own rise from 3.0% to 5.2% agrees with Diop (2018), who concludes that non-spatial estimates may be considerably understated. Others find the opposite: lower speeds for European regions (Isla-Castillo et al. 2024), neighbor effects that slow convergence across Indonesian districts (Santos-Marquez et al. 2022), and divergence rather than convergence across Indian states (Lolayekar, Mukhopadhyay 2019). The substantive point is therefore not that spatial dependence matters, which is already known, but that here it works in the direction of faster measured convergence, and that the scale of analysis is part of what determines the answer.¹⁵

¹⁴The regularity is about 2% per year across 1,528 regions in 83 countries (Gennaioli et al. 2014), and similar across regions of the United States, Japan, and five European countries (Sala-i-Martin 1996). The meta-analytic objection is that estimates from different specifications are drawn from different populations, so there is no single natural rate to be estimated.

¹⁵In our own case the indirect effect carries the same negative sign as the direct effect, so the two reinforce one another and the total effect exceeds the OLS coefficient in magnitude. Where the initial conditions of neighbors enter with the opposite sign, or where the spatial terms absorb variation the non-spatial model had attributed to initial luminosity, the same exercise reduces the estimated speed.

3.5 Robustness to alternative spatial weight matrices

The spillover results above use the 6NN spatial weight matrix defined in Section 2. To assess their sensitivity to this choice, we re-estimate the preferred Model 4 under six alternative specifications and compare them with the 6NN baseline. The alternatives are four- and eight-nearest neighbors, queen and rook contiguity, and inverse distance and inverse-distance-squared, the last two applied within a distance band.



Notes: Points are Direct, Indirect, and Total impacts; bars are 95% Monte-Carlo confidence intervals. The dashed lines mark the 6NN baseline estimates.

Figure 8: Robustness of the Model 4 spatial impacts of initial luminosity to the choice of spatial weight matrix

Table 4: Model 4 spatial impacts of initial luminosity under alternative spatial weight matrices

Weight matrix	Direct	Indirect	Total	AIC
4 nearest neighbors	-0.024*** (0.002)	-0.008 (0.005)	-0.032*** (0.006)	-2468
6 nearest neighbors (base)	-0.025*** (0.002)	-0.012* (0.007)	-0.037*** (0.007)	-2501
8 nearest neighbors	-0.025*** (0.002)	-0.011 (0.008)	-0.036*** (0.008)	-2485
Queen contiguity	-0.025*** (0.002)	-0.010** (0.005)	-0.035*** (0.005)	-2463
Rook contiguity	-0.025*** (0.002)	-0.009** (0.004)	-0.034*** (0.005)	-2469
Inverse distance	-0.025*** (0.002)	-0.016** (0.007)	-0.041*** (0.008)	-2486
Inverse distance squared	-0.025*** (0.002)	-0.012* (0.006)	-0.037*** (0.006)	-2485

Notes: Full LeSage–Pace method; Monte-Carlo standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The direct and total convergence effects are robust to the choice of spatial weights, and the spillover component is consistently negative but more sensitive (Figure 8 and Table 4). The direct (within-district) convergence effect is essentially unchanged across all seven specifications, ranging from -0.024 to -0.025 and always significant at the 1% level. The total convergence effect likewise remains negative and significant throughout, ranging from -0.032 to -0.041, with the 6NN baseline (-0.037) near the middle of this range. The indirect (spillover) component is negative in every specification and statistically significant under five of the seven, which is a consistent if weaker pattern. This stability indicates that the direct and total convergence effects are not an artifact of the particular neighbor definition.

The results reported so far treat luminosity as a proxy for economic output, which is the use that the convergence literature has made of it. The discussion that follows asks first

what the magnitude of these estimates implies for regional policy, and what rests on the use of a single growth period. It then turns to a second use of the same data, since nighttime lights also record electrification, urbanization, and broader socioeconomic conditions, and the reproducible workflow assembled here can be pointed at those dimensions as readily as at output. That possibility is pursued in an exploratory way, asking how luminosity relates to cultural participation across Indian states.

4 Discussion

4.1 *What the magnitude of the effects implies for regional policy*

The difference between the two readings of Model 4 is large enough to matter for policy. Under the OLS estimate, a district closes half of the gap to its steady state in about 23 years; under the spatial Durbin estimate, in about 13 years. Over the fourteen years of our sample, that is the difference between closing roughly 34% and roughly 52% of the initial gap. The first figure describes catch-up over a generation; the second, catch-up within the horizon over which development programs are designed, funded, and evaluated. Under the faster rate, the returns to well-targeted support should appear within a normal evaluation window.

The results also speak to the level at which such policies should operate. We find convergence across districts over 1996–2010, whereas [Lolayekar, Mukhopadhyay \(2019\)](#), working with Indian states over a longer period, find divergence. Catch-up visible at the district level and absent at the state level is consistent with the district being the more informative unit of intervention, which the Aspirational Districts Programme of India already assumes.¹⁶ The spatial results add a qualification, since a lagging district surrounded by other lagging districts is in a different position from one adjacent to a growing cluster. A targeting rule that ranks districts one at a time will not register that difference.

Spillovers cut both ways, however, and the Indian evidence on place-based policy is a caution as much as an encouragement. [Hasan et al. \(2021\)](#) find that a tax exemption for industrially backward districts raised firm entry and employment partly by displacing activity from neighboring districts that narrowly missed qualifying, with effects that did not outlast the program. Displacement of that kind would appear in our framework as a spatial relationship between neighbors, which is not a spillover on which coordinated regional policy can be built. What the satellite record does support is measurement.¹⁷ The readings above therefore hold only if the indirect effects reflect genuine interaction between districts, and the reduced-form evidence is consistent with that but does not establish it.

4.2 *The single growth period and what a panel would require*

The estimates reported above rest on a single growth period, and the cost of that design should be stated plainly. One long difference cannot trace how convergence evolved within the period, cannot show whether the spatial spillovers were stronger in some years than in others, and cannot separate long-run convergence from what was specific to these fourteen years, which in India included the maturing of the post-1991 reforms. Examining those dynamics lies beyond an article that adds a spatial dimension to an existing cross-sectional result. The immediate obstacle is in any case the data, since district population is observed only at the decennial censuses.¹⁸ Interpolated denominators touch only the endpoints of

¹⁶The program was launched in 2018 to direct additional resources to the least developed districts of the country. The two studies also differ in outcome variable and period, so the contrast between them is suggestive rather than decisive.

¹⁷[Shenoy \(2018\)](#) uses nighttime lights to evaluate a large place-based package in Uttarakhand and finds that luminosity rises sharply on the targeted side of the new state border, which shows that the outcome we study is responsive enough to register policy at this scale.

¹⁸The dataset we adopt from [Chanda, Kabiraj \(2020\)](#) holds radiance-calibrated luminosity at six dates between 1996 and 2010, so per-capita luminosity at the intermediate dates would rest on interpolated denominators.

a long difference, whereas in a panel they would enter every observation, and would do so in the time dimension that a panel exists to exploit.

What such a panel would be worth is the more interesting question, and here two problems compound. The first predates nighttime lights: panel estimates of convergence run far above cross-sectional ones, and [Caselli et al. \(1996\)](#) obtain a rate near 10% per year against the 2% that cross-sections deliver.¹⁹ How that gap should be read remains contested, so a panel would give not a better-identified version of our estimate but an estimate of something different. The second problem is specific to the proxy, since luminosity discriminates differences across places far better than changes over time ([Chen, Nordhaus 2019](#), [Asher et al. 2021](#)), and over short intervals the ratio of signal to measurement error falls sharply.²⁰ The periods compared should therefore be long, and harmonized DMSP–VIIRS series now make a sequence of non-overlapping long differences feasible across roughly three decades ([Li et al. 2020](#)), which is the most promising temporal extension of this work.

4.3 An exploratory extension: Luminosity and cultural participation

Nighttime lights capture not just GDP but a composite of electrification, urbanization, infrastructure, and broader socioeconomic conditions ([Henderson et al. 2012](#), [Mellander et al. 2015](#)). In this section, we examine the association between luminosity and cultural participation patterns across Indian states.

A growing literature argues that regional cultural attitudes constitute an independent factor in economic development, not merely a by-product of income or institutions. In particular, [Tubadji \(2025\)](#) applies the Culture-Based Development framework, whose central claim is a directional one. Cultural attitudes and the participation patterns that reveal them are treated as a factor shaping local development, rather than as an outcome that development produces. That direction also marks the limit of what the exercise below can establish.²¹

For India, the National Sample Survey (NSS) 47th Round (July–December 1991) surveys household cultural participation, and nighttime lights for the adjacent year 1992 are available for 32 states and union territories. Following the Culture-Based Development framework, we reduce these survey items to six dimensions and aggregate them to state means.²² This part of the analysis uses the Consistent and Corrected Nighttime Lights series rather than the radiance-calibrated composites used above, because it is the product that covers 1992. Given the small sample ($N = 32$) and the leverage that small territories at the extremes of the distribution exert on linear correlation measures, we report Spearman rank correlations throughout. Figure 9 presents the relationship between log nighttime lights per capita and the two cultural dimensions that show statistically significant associations.

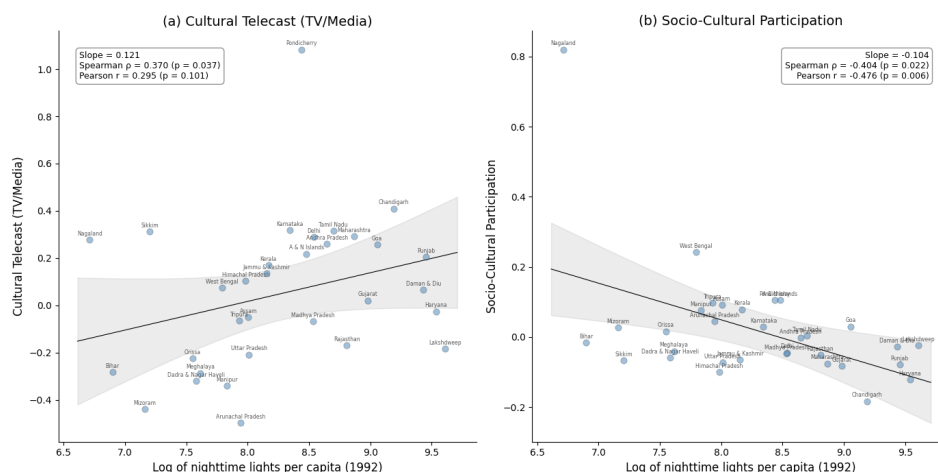
Cultural telecast (TV/media) is positively associated with luminosity, with a Spearman rank correlation of 0.370 ($p = 0.037$), so more luminous states tend to rank higher on that dimension. Socio-cultural participation, in contrast, is negatively associated, with a Spearman rank correlation of -0.404 ($p = 0.022$), so less luminous states tend to rank higher on it. We refer to these two dimensions as media-based cultural consumption and community-based participation, respectively. This suggests that less urbanized regions rely more on collective, in-person forms of cultural participation than on media infrastructure.

¹⁹[Barro \(2015\)](#) argues that country fixed effects generate a misleadingly high convergence rate, and that combining post-1960 with post-1870 evidence returns a figure close to the conventional 2%. Meta-analytic evidence finds that corrections for unobserved heterogeneity raise estimated convergence systematically ([Abreu et al. 2005](#)).

²⁰A short-horizon panel would therefore pair a specification that is hard to interpret with a proxy operating where it is weakest. [Chakrabarty, Mukherjee \(2022\)](#) estimate district convergence across the pandemic period, and how such estimates should be read depends on how much of the short-run movement in lights reflects output rather than measurement.

²¹A cross-sectional association is consistent with either direction, and with both operating at once. What it can do is establish that the two are related, and describe the form that relationship takes.

²²The six dimensions are live cultural performance, cultural telecast (TV/media), socio-cultural participation, cultural heritage and religion, live cultural shows, and sports. They are extracted from the survey items by principal-components analysis. See the [Spatial culture](#) notebook for details and extended analyses.



Source: Nighttime lights from CCNL DMSP-OLS (Zhao et al., 2022). Cultural participation from NSS 47th Round (July–December 1991). See [Spatial culture](#) notebook for source code.

Notes: Each point represents one of the 32 Indian states and union territories. Panel (a) shows the bivariate relationship between log nighttime lights per capita and cultural telecast (TV/media); Panel (b) shows the relationship with socio-cultural participation. Solid line is the OLS regression; gray band shows the 95% confidence interval (t-distribution, 30 df). Annotations report regression slope, Spearman rank correlation, and Pearson correlation.

Figure 9: Relationship between nightlight luminosity and cultural participation across 32 Indian states

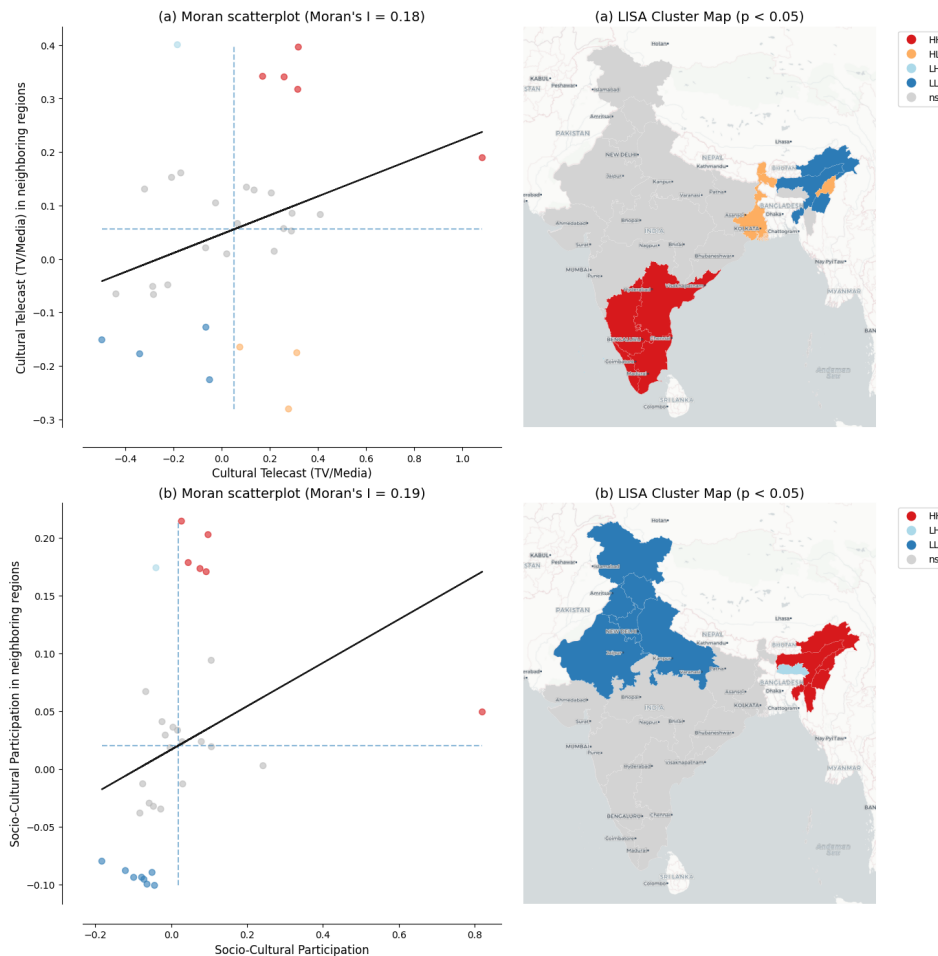
Figure 10 maps the spatial distribution of the two cultural variables through the lens of a LISA analysis. The Low-Low cluster of media-based cultural consumption in the northeast coincides with the Low-Low cluster in the nighttime lights data, whereas the High-High clusters do not coincide: media-based cultural consumption forms one in the southern states, and community-based participation forms one in the northeast. The remaining four cultural dimensions show no statistically significant association with luminosity, and these findings rest on a small cross-section of 32 states observed at a single point in time, so larger datasets covering more regions and periods would be needed to establish whether the patterns are robust.

Two connections to the main analysis are worth drawing out, both offered as interpretation rather than as findings. The first concerns what luminosity measures. A positive association with media-based cultural consumption and a negative one with community-based participation are what one would expect if luminosity tracks electrification and urbanization rather than economic output alone. The cultural results are, in that sense, a small piece of evidence about the composite nature of the proxy on which the convergence analysis rests, and they bear on the measurement issues discussed below.

The second connection concerns spatial structure. Cultural participation is not distributed randomly across states, since both dimensions examined here are significantly spatially autocorrelated. Factors of this kind, spatially clustered and correlated with development, are precisely what a spatial Durbin model absorbs into its indirect effects when they are not measured directly. We cannot test this with state-level data for a single adjacent period, and we do not claim that cultural participation drives the spillovers estimated in the results section. What the exercise does is make concrete what a reduced-form spatial model may be capturing, and it suggests that the Culture-Based Development framework offers one route to testing the question directly, should comparable data become available at the district level and over time.

4.4 New research directions

Three kinds of gap limit what this literature can currently establish: the data themselves, the geography they cover, and the theories they have been used to test. The first is well documented, since the radiance-calibrated DMSP-OLS series we adopt from [Chanda, Kabiraj \(2020\)](#), and on which this literature rests ([Henderson et al. 2012](#), [Chen, Nordhaus](#)



Source: Cultural participation data from NSS 47th Round (July–December 1991). See [Spatial culture](#) notebook for source code.

Notes: Panel (a) shows results for cultural telecast (TV/media); Panel (b) for socio-cultural participation. Left subpanels show Moran scatterplots with the Global Moran's I statistic. Right subpanels show LISA cluster maps with statistically significant clusters at $p < 0.05$ based on 999 permutations and a 6-nearest-neighbors spatial weights matrix. Region labels are overlaid from the CartoDB Positron basemap.

Figure 10: LISA cluster maps of cultural participation across 32 Indian states

2011), is top-coded in bright urban cores, uncalibrated on board, and blurred across boundaries by blooming (Gibson et al. 2021, Abrahams et al. 2018). Top-coding leaves the net bias in our convergence coefficient ambiguous in sign, while blooming raises the apparent correlation between neighbors and therefore bears directly on the spillover estimates.²³ The Visible Infrared Imaging Radiometer Suite (VIIRS) and harmonized DMSP–VIIRS series address all three problems (Elvidge et al. 2017, Li et al. 2020, Chiovelli et al. 2026), and multi-source methods improve measurement where lights alone are weak.²⁴ Re-estimating our specification on those data is the most immediate extension

²³Top-coding compresses variation at the bright end of the distribution, so classical measurement error in initial luminosity attenuates the convergence coefficient toward zero, while the same error entering the growth rate with the opposite sign works in the direction of overstating convergence. The net effect is ambiguous in sign, and we do not claim to sign it. Measurement error is also unlikely to be homogeneous across the sample, because the association between luminosity and economic outcomes is stronger in urban than in rural areas and Indian districts differ considerably in how urbanized they are. That the direct and total effects hold across all seven definitions of the neighborhood, and the indirect effect under five of the seven (Table 4), indicates that they are not an artifact of any particular neighbor definition, but it does not rule out blooming, for which deblurring methods offer one route (Abrahams et al. 2018).

²⁴VIIRS, operational since 2012, improves spatial resolution (roughly 750 meters against 2.7 kilometers), adds on-board radiometric calibration, and widens dynamic range, which avoids saturation in bright areas while detecting dim lights in rural settlements (Gibson et al. 2021). Lights have also been combined

of the analysis reported here.

The second gap concerns geography, and it is not that some regions have gone unstudied. A single instrument observes the whole planet under a common protocol, which makes measurements comparable in a way that national accounts are not (Henderson et al. 2012), but what it records is light rather than economic activity. Roughly one fifth of the settlement footprint of the planet emits no detectable radiance, and unlit settlements are concentrated in Africa (McCallum et al. 2022). Africa is nonetheless well covered by this literature (Kacou 2022), as are China (Glawe, Mendez 2025), Indonesia (Miranti, Mendez 2023), Thailand (Tipayalai, Mendez 2024), and Turkey (Ursavaş, Mendez 2023). The real gap is that the places where lights carry least information, rural and sparsely electrified areas, are the places where administrative statistics are weakest and a satellite proxy would be most valuable.²⁵ Closing it calls less for further applications of the same proxy than for validation against ground data in those settings.

The third gap concerns theory, since luminosity is now applied well beyond the question asked here. Within convergence analysis, our framework yields an average effect rather than heterogeneity across the distribution, and it captures spillovers in reduced form, gaps that distribution dynamics and quasi-experimental designs such as that of Asher, Novosad (2020) are built to fill.²⁶ Beyond convergence, lights now serve the study of city size (Ch et al. 2021), electrification and the provision of public goods (Min et al. 2013), armed conflict (Witmer, O’Loughlin 2011), the cultural participation theorized by Tubadji (2025), and the local effect of policies and shocks at fine spatial scales (Bluhm, McCord 2022). One caveat governs all of these uses: lights discriminate differences across places far better than changes over time, so evaluations over short horizons require validation against local ground data.²⁷ The design used here, a fourteen-year horizon, rests mostly on cross-sectional variation, which is what lights currently support best.

Taken together, these three gaps describe a research agenda rather than a list of caveats, and none of the work they imply requires starting from scratch. The pipeline behind this article is public, documented, and executable in a browser, so a reader can open the relevant notebook, substitute data from another country or a later period, and re-run it. That is the contribution we would most like this article to make: not an estimate of how quickly Indian districts converged between 1996 and 2010 alone, but a working template that lowers the cost of asking the same question elsewhere.

4.5 Research reproducibility and open science

As Donaldson, Storeygard (2016) note, methodological choices in processing remotely sensed data, such as sensor inter-calibration, can affect the empirical conclusions. The same applies further down our own pipeline, at the construction of the spatial weight matrix. Our own robustness analysis is a case in point: the estimated convergence effects survive seven definitions of the neighborhood, but that is something a reader can only confirm because the code that produced each of them is available.

Open-source tools and cloud platforms lower the barriers to that kind of verification,

with daytime imagery for poverty prediction (Jean et al. 2016), with land cover in agricultural areas (Keola et al. 2015), with multiple remote sensing indicators for subnational output (Chen et al. 2024, Suleiman et al. 2025), and with survey data for multidimensional poverty (Khoun et al. 2025).

²⁵Africa accounts for 39% of unlit settlements worldwide, rising to 65% when only rural settlement areas are considered (McCallum et al. 2022). Across 34 sub-Saharan African countries the association between harmonized luminosity and household wealth is pronounced in urban areas and considerably weaker in rural ones (Jhamb et al. 2025), and systematic assessments of which product to use, and where, reach similar conclusions (Gibson et al. 2021).

²⁶The relevant tools are distribution dynamics (Quah 1996) and the regression-tree methods developed for identifying multiple growth regimes (Durlauf, Johnson 1995), which can reveal whether Indian districts converge to a single steady state or form distinct clubs. Combining them with spatial econometrics could expose spatial poverty traps or growth poles that average estimates render invisible (Rey, Montouri 1999).

²⁷VIIRS predicts cross-sectional GDP better than time-series GDP (Chen, Nordhaus 2019), and in India the elasticity of lights with respect to local outcomes is far lower in the time series than in the cross-section (Asher et al. 2021). The relationship between lights and growth is also unstable across regions in both Brazil and India (Bickenbach et al. 2016), and responsiveness to output varies with baseline income, density, and sector. An evaluation may therefore fail to detect a real effect where lights react weakly, or attribute to a policy what is in fact heterogeneity in the relationship between light and output.

and cloud-based notebooks for processing nighttime lights remove the need for specialized local software (Patnaik, Mendez 2024). Open data, version-controlled code, and cloud execution together make the full pipeline, from satellite imagery to econometric results, verifiable and extensible by the broader research community.

5 Concluding remarks

This article re-examines regional convergence across Indian districts using satellite nighttime lights, interactive visualization, and spatial econometric modeling. Our tests show that spatial dependence characterizes both the satellite data and the convergence process itself. In the fully specified spatial Durbin model, the total convergence effect is about 51% larger than the conventional non-spatial estimate, and the implied half-life of regional disparities falls from roughly 23 years to 13. Non-spatial models therefore understate the speed of convergence in this setting, although comparison with other regional literatures shows that the direction of the adjustment is specific to the setting. The policy reading depends on what the indirect effects represent. Genuine spillovers would mean that place-based interventions reach beyond the districts they target. Displaced activity would mean that the same estimates overstate what those interventions add in aggregate.

Every step of the analysis is documented in Python notebooks, which run in the cloud and are embedded in the manuscript through the Quarto publishing framework.²⁸ The code and the data are hosted in a public GitHub repository, and we follow these open-science practices in the hope that other applied studies will adopt them. That same workflow allowed us to extend the analysis beyond economic output. Relating luminosity to cultural participation across Indian states required one additional notebook. Luminosity correlates positively with media-based cultural consumption and negatively with community-based participation. Those exploratory findings are reproducible on the same terms as the convergence estimates.

The most valuable next steps are concrete: re-estimate this specification on the higher-resolution VIIRS data, validate luminosity against ground measurements in rural and sparsely electrified areas, and carry the same spatial framework to questions beyond convergence. The convergence estimates reported here will eventually be superseded, and the workflow that produced them is built to be reused.

Acknowledgments

This research project was supported by JSPS KAKENHI Grant Number 24K04884. During the preparation of this manuscript, the authors acknowledge the use of Claude Code (Anthropic) to assist with manuscript editing, computational notebook development, and research infrastructure setup. After using this AI tool, the authors reviewed the outputs, confirmed their accuracy, and take full responsibility for the content of this publication.

Conflict of interest

The authors declare no conflict of interest.

Data and code availability

All data and computational code used in this study are available in the project repository at <https://github.com/quarcs-lab/project2025s-py>. The interactive HTML version of this manuscript (<https://quarcs-lab.github.io/project2025s-py/>) embeds the computational notebooks, allowing readers to inspect the complete analytical pipeline from raw data to published results. Every analytical notebook can be executed in the cloud using Google Colaboratory without any local software installation: readers open the notebook and

²⁸A single manuscript source generates multiple output formats. The HTML version allows readers to engage with interactive visualizations, inspect the code of the computational notebooks, and re-evaluate the results in light of their source code.

run it, and the empirical results, figures, and tables of this article are regenerated from the raw data. The computational notebooks are implemented entirely in Python. An earlier implementation of the same analysis used R and Stata, and the present Python results reproduce it. The AIC values differ by a few units because Stata and Python count parameters differently: Stata treats the error variance of the spatial model as a free parameter, and its information criterion for the fixed-effects regressions uses the rank of the robust covariance matrix, which two single-district states leave short by two. The interactive Earth Engine application linked in the text is a convenience layer for visualizing the underlying imagery; its source code is included in the first notebook, and none of the results reported in this article depend on it.

References

- Abrahams A, Oram C, Lozano-Gracia N (2018) Deblurring DMSP nighttime lights: A new method using Gaussian filters and frequencies of illumination. *Remote Sensing of Environment* 210: 242–258. [CrossRef](#)
- Abreu M, de Groot HLF, Florax RJGM (2005) A meta-analysis of β -convergence: The legendary 2%. *Journal of Economic Surveys* 19: 389–420. [CrossRef](#)
- Adhikari B, Dhital S (2021) Decentralization and regional convergence: Evidence from night-time lights data. *Economic Inquiry* 59: 1066–1088. [CrossRef](#)
- Allaire JJ, Teague C, Scheidegger C, Xie Y, Dervieux C (2024). Quarto. <https://quarto.org>
- Anselin L (1988) *Spatial Econometrics: Methods and Models*. Springer. [CrossRef](#)
- Anselin L (1995) Local indicators of spatial association—LISA. *Geographical Analysis* 27: 93–115. [CrossRef](#)
- Asher S, Lunt T, Matsuura R, Novosad P (2021) Development research at high geographic resolution: An analysis of night-lights, firms, and poverty in India using the SHRUG open data platform. *The World Bank Economic Review* 35: 845–871. [CrossRef](#)
- Asher S, Novosad P (2020) Rural roads and local economic development. *American Economic Review* 110: 797–823. [CrossRef](#)
- Barro RJ (2015) Convergence and modernisation. *The Economic Journal* 125: 911–942. [CrossRef](#)
- Barro RJ, Sala-i-Martin X (1992) Convergence. *Journal of Political Economy* 100: 223–251. [CrossRef](#)
- Basher SA, Rashid S, Uddin MR (2023) Regional convergence in Bangladesh using night lights. *Applied Economics Letters* 30: 2581–2588. [CrossRef](#)
- Beyer RCM, Jain T, Sinha S (2023) Lights out? COVID-19 containment policies and economic activity. *Journal of Asian Economics* 85: 101589. [CrossRef](#)
- Beyer RCM, Yamada T (2020) District-level convergence and public banks in India. SSRN Working Paper 3818578, Social Science Research Network. [CrossRef](#)
- Bickenbach F, Bode E, Nunnenkamp P, Söder M (2016) Night lights and regional GDP. *Review of World Economics* 152: 425–447. [CrossRef](#)
- Bluhm R, McCord GC (2022) What can we learn from nighttime lights for small geographies? measurement errors and heterogeneous elasticities. *Remote Sensing* 14: 1190. [CrossRef](#)
- Carrington SJ, Jiménez-Ayora P (2021) Shedding light on the convergence debate: Using luminosity data to investigate economic convergence in Ecuador. *Review of Development Economics* 25: 200–227. [CrossRef](#)
- Caselli F, Esquivel G, Lefort F (1996) Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth* 1: 363–389. [CrossRef](#)

- Ch R, Martin DA, Vargas JF (2021) Measuring the size and growth of cities using nighttime light. *Journal of Urban Economics* 125: 103254. [CrossRef](#)
- Chakrabarty M, Mukherjee S (2022) Impact of COVID-19 on convergence in Indian districts. *Indian Growth and Development Review* 15: 198–209. [CrossRef](#)
- Chakravarty P, Dehejia V (2017) Will GST exacerbate regional divergence? *Economic and Political Weekly* 52: 97–102. <https://www.epw.in/journal/2017/25-26/notes/will-gst-exacerbate-regional-divergence.html>
- Chanda A, Cook CJ (2022) Was India's demonetization redistributive? Insights from satellites and surveys. *Journal of Macroeconomics* 73: 103438. [CrossRef](#)
- Chanda A, Kabiraj S (2020) Shedding light on regional growth and convergence in India. *World Development* 133: 104961. [CrossRef](#)
- Chen X, Nordhaus WD (2011) Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences* 108: 8589–8594. [CrossRef](#)
- Chen X, Nordhaus WD (2019) VIIRS nighttime lights in the estimation of cross-sectional and time-series GDP. *Remote Sensing* 11: 1057. [CrossRef](#)
- Chen Y, Ursavaş U, Mendez C (2024) Can higher-quality nighttime lights predict sectoral GDP across subnational regions? Urban and rural luminosity across provinces in Türkiye. *Letters in Spatial and Resource Sciences* 17: 12. [CrossRef](#)
- Chiovelli G, Michalopoulos S, Papaioannou E, Regan T (2026) Illuminating the Global South. *The Economic Journal*. [CrossRef](#)
- Cook CJ, Shah M (2022) Aggregate effects from public works: Evidence from India. *Review of Economics and Statistics* 104: 797–806. [CrossRef](#)
- Diop S (2018) Convergence and spillover effects in Africa: A spatial panel data approach. *Journal of African Economies* 27: 274–284. [CrossRef](#)
- Donaldson D, Storeygard A (2016) The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives* 30: 171–198. [CrossRef](#)
- Durlauf SN, Johnson PA (1995) Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10: 365–384. [CrossRef](#)
- Elvidge CD, Baugh KE, Zhizhin M, Hsu FC, Ghosh T (2017) VIIRS night-time lights. *International Journal of Remote Sensing* 38: 5860–5879. [CrossRef](#)
- Ertur C, Koch W (2007) Growth, technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics* 22: 1033–1062. [CrossRef](#)
- Fischer MM (2011) A spatial Mankiw-Romer-Weil model: Theory and evidence. *Annals of Regional Science* 47: 419–436. [CrossRef](#)
- Gennaioli N, La Porta R, Lopez-de-Silanes F, Shleifer A (2014) Growth in regions. *Journal of Economic Growth* 19: 259–309. [CrossRef](#)
- Gibson J, Olivia S, Boe-Gibson G, Li C (2021) Which night lights data should we use in economics, and where? *Journal of Development Economics* 149: 102602. [CrossRef](#)
- Glawe L, Mendez C (2025) Harmonized luminosity and economic activity across provinces in China: Cross-sectional differences, regional time series, and inequality dynamics. *Applied Economics* 57: 10677–10693. [CrossRef](#)
- Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R (2017) Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202: 18–27. [CrossRef](#)
- Hasan R, Jiang Y, Rafols RM (2021) Place-based preferential tax policy and industrial development: Evidence from India's program on industrially backward districts. *Journal of Development Economics* 150: 102621. [CrossRef](#)

- Henderson JV, Storeygard A, Weil DN (2012) Measuring economic growth from outer space. *American Economic Review* 102: 994–1028. [CrossRef](#)
- Isla-Castillo F, Garashchuk A, Podadera-Rivera P (2024) Cross-sectional and spatial panel data analysis of territorial economic cohesion in the European Union regions based on convergence approach: From 2 to 8 per cent? *Socio-Economic Planning Sciences* 95: 102012. [CrossRef](#)
- Jean N, Burke M, Xie M, Davis WM, Lobell DB, Ermon S (2016) Combining satellite imagery and machine learning to predict poverty. *Science* 353: 790–794. [CrossRef](#)
- Jha P, Talathi K (2024) Impact of colonial institutions on economic growth and development in India: Evidence from night-lights data. *Economic Development and Cultural Change* 72: 1653–1708. [CrossRef](#)
- Jhamb P, Ferreira S, Stephens P, Sundaram M, Wilson J (2025) Shedding light on development: Leveraging the new nightlights data to measure economic progress. *PLOS ONE* 20: e0318482. [CrossRef](#)
- Kacou KYT (2022) Interregional inequality in Africa, convergence, and multiple equilibria: Evidence from nighttime light data. *Review of Development Economics* 26: 918–940. [CrossRef](#)
- Keola S, Andersson M, Hall O (2015) Monitoring economic development from space: Using nighttime light and land cover data to measure economic growth. *World Development* 66: 322–334. [CrossRef](#)
- Khoun T, Poortinga A, Thwal NS, González de Alba I, McMahon A, Mendez C (2025) Mapping the dimensions of poverty through big data, socioeconomic surveys and machine learning in Cambodia. *Social Indicators Research* 180: 1593–1618. [CrossRef](#)
- Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Kelley K, Hamrick J, Grout J, Corlay S et al. (2016) Jupyter notebooks—a publishing format for reproducible computational workflows. In: Schmidt B, Loizides F (eds), *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. IOS Press, Amsterdam, NL, 87–90
- Knuth DE (1984) Literate programming. *The Computer Journal* 27: 97–111. [CrossRef](#)
- LeSage JP, Pace RK (2009) *Introduction to Spatial Econometrics*. Chapman and Hall/CRC. [CrossRef](#)
- Lessmann C, Seidel A (2017) Regional inequality, convergence, and its determinants: A view from outer space. *European Economic Review* 92: 110–132. [CrossRef](#)
- Li X, Zhou Y, Zhao M, Zhao X (2020) A harmonized global nighttime light dataset 1992–2018. *Scientific Data* 7: 168. [CrossRef](#)
- Lolayekar AP, Mukhopadhyay P (2019) Spatial dependence and regional income convergence in India (1981–2010). *GeoJournal* 84: 851–864. [CrossRef](#)
- McCallum I, Kyba CCM, Bayas JCL, Moltchanova E, Cooper M, Cuaresma JC, Pachauri S, See L, Danylo O, Moorthy I, Lesiv M, Baugh K, Elvidge CD, Hofer M, Fritz S (2022) Estimating global economic well-being with unlit settlements. *Nature Communications* 13: 2459. [CrossRef](#)
- Mellander C, Lobo J, Stolarick K, Matheson Z (2015) Night-time light data: A good proxy measure for economic activity? *PLOS ONE* 10: e0139779. [CrossRef](#)
- Min B, Gaba KM, Sarr OF, Agalassou A (2013) Detection of rural electrification in Africa using DMSP-OLS night lights imagery. *International Journal of Remote Sensing* 34: 8118–8141. [CrossRef](#)
- Miranti RC, Mendez C (2023) Social and economic convergence across districts in Indonesia: A spatial econometric approach. *Bulletin of Indonesian Economic Studies* 59: 421–445. [CrossRef](#)

- Moran PAP (1950) Notes on continuous stochastic phenomena. *Biometrika* 37: 17–23. [CrossRef](#)
- Patnaik A, Mendez C (2024) Exploring economic activity from outer space: A Python notebook for processing and analyzing satellite nighttime lights. *REGION* 11: 79–109. [CrossRef](#)
- Peng RD (2011) Reproducible research in computational science. *Science* 334: 1226–1227. [CrossRef](#)
- Pinkovskiy M, Sala-i-Martin X (2016) Lights, camera, income! illuminating the national accounts-household surveys debate. *Quarterly Journal of Economics* 131: 579–631. [CrossRef](#)
- Quah DT (1996) Twin peaks: Growth and convergence in models of distribution dynamics. *The Economic Journal* 106: 1045–1055. [CrossRef](#)
- Rey SJ, Montouri BD (1999) US regional income convergence: A spatial econometric perspective. *Regional Studies* 33: 143–156. [CrossRef](#)
- Sala-i-Martin X (1996) Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review* 40: 1325–1352. [CrossRef](#)
- Santos-Marquez F, Gunawan AB, Mendez C (2022) Regional income disparities, distributional convergence, and spatial effects: Evidence from Indonesian regions 2010–2017. *GeoJournal* 87: 2373–2391. [CrossRef](#)
- Sanyal S, Arora A (2024) Relative economic performance of Indian states: 1960-61 to 2023-24. Working Paper EAC-PM/WP/31/2024, Economic Advisory Council to the Prime Minister
- Shenoy A (2018) Regional development through place-based policies: Evidence from a spatial discontinuity. *Journal of Development Economics* 130: 173–189. [CrossRef](#)
- Solow RM (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70: 65–94. [CrossRef](#)
- Suleiman H, Nguyen MTT, Mendez C (2025) Predicting subnational GDP in Vietnam with remote sensing data: A machine learning approach. *Letters in Spatial and Resource Sciences* 18: 5. [CrossRef](#)
- Tamiminia H, Salehi B, Mahdianpari M, Quackenbush LJ, Adeli S, Brisco B (2020) Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing* 164: 152–170. [CrossRef](#)
- Tipayalai K, Mendez C (2024) Regional convergence and spatial dependence in Thailand: Global and local assessments. *Journal of the Asia Pacific Economy* 29: 693–720. [CrossRef](#)
- Tubadji A (2025) Cultural entropy, innovation, and growth. *Politics & Policy* 53: e70050. [CrossRef](#)
- Ursavaş U, Mendez C (2023) Regional income convergence and conditioning factors in Turkey: Revisiting the role of spatial dependence and neighbor effects. *Annals of Regional Science* 71: 363–389. [CrossRef](#)
- Witmer FDW, O'Loughlin J (2011) Detecting the effects of wars in the Caucasus regions of Russia and Georgia using radiometrically normalized DMSP-OLS nighttime lights imagery. *GIScience & Remote Sensing* 48: 478–500. [CrossRef](#)

