Indian Public Policy Review 2022, 3(5): 49-86
https://doi.org/10.55763/ippr.2022.03.05.003

IPPR

Conditional convergence and Spatial convergence across 103 Sub-state Indian regions:
Using spatial econometrics for panel data

Vivek Jadhav[*] **Brinda Viswanathan**[#]

Abstract

Recent studies have devoted great emphasis to examining the phenomenon of income convergence across regions. The empirical efforts made in the context of India look at convergence among the states of India. Although there have been a few studies done on the district level, the sub-state regions that are prevalent within each state have been largely ignored in the Indian regional literature. The purpose of this research is to investigate the extent to which 103 sub-state regions within 20 Indian states converge. This research adopts a method that differs from the conventional convergence strategy by instead focusing on the spatial convergence aspect. It has been shown that not only does spatial convergence but also β-convergence: a growth process where poor regions grow faster than rich regions occur among India's 103 different regions. This study sheds insight on the two distinct forms of convergence, namely, β-convergence across all regions, and β-convergence among neighbouring regions. The finding of the existence of β-convergence and spatial convergence among neighbouring regions invites policy attention regarding the development of backward regions.

Keywords: Spatial convergence, Spatial fixed effect models, Panel data, Conditional Convergence.

JEL Classification: C31, R12, C33, O41

Publication Date: 23 September 2022

[*] Vivek Jadhav is pursuing a Ph.D. in Economics from Madras School of Economics.
[#] Dr. Brinda Viswanathan is Professor, Madras School of Economics, Chennai

1. Introduction

India's per capita GDP has climbed from USD 758 in 2000 to USD 1606 in 2015. Because of this impressive national growth, attention has been drawn to the regional growth that occurred during the same period. It's interesting to see if the growth of regions is at par with national growth.

In the context of the regional economy, growth has always attracted debate about convergence and divergence. The concept of convergence is not new since it was proposed by Robert Solow's growth model. In literature, two broad concepts of convergence are discussed[1]: β-convergence and σ-convergence. β-convergence reflects to a process where poor economies grow faster than rich economies and σ-convergence reflects the process where the differences in the real GDP of economies tends to decrease. Neoclassical exogenous growth theory and augmented growth theory (Solow-Swan, 1956; Mankiw, Romer, and Weil (MRW), 1992; Barro and Sala-I-Martin, 1995) believe that early per capita income differences will ultimately be 'conditionally converged[2]' due to capital accumulation and diminishing returns. This convergence argument is also backed by region as well as country-specific recent empirical studies (Maddison, 1991; Barro and Sala-i-Martin, 1991, 1992; Cashin, 1995; Sala-i-Martin, 1996; Armstrong, 1995; Persson, 1995; Cashin and Sahay, 1996; Barro-Lee, 2001 data set; Singh, et al, 2010).

In the context of India, the literature on convergence is vast and expanding. The recent development in the convergence literature also incorporates spatial aspects (Shaban, 2006; Kocornik-Mina, 2009; Kalra and Thakur, 2015). While empirical studies incorporate recent theoretical advances, they are limited to the state level. This study shifts the level of analysis for regional studies from Indian states to sub-state regions. The regions in this research are distinct from 'Indian states'. The regions considered in this study are administrative and geopolitical divisions within its states, comprising districts. This research first defines the regions in India and then tries to understand the convergence as well as spatial convergence using panel data for the 2001-2015 period.

The regions can be described in many ways depending on their characteristics. A region is characterised primarily by its size, content, location, and border. The region also has another characteristic, and that is homogeneity (Malgavkar & Ghiara, 1969). A set of countries, states, districts, or villages might be referred to as a region.

When applied to the context of India, regional studies tend to focus more on the Indian states in convergence literature. In the context of this study, a region refers to a set of districts that have similar characteristics and come together to create administrative divisions. In accordance with the concept presented by Malgavkar & Ghiara (1969), these regions are not only administrative divisions, but they are also homogeneous in terms of social identity, which includes religion and caste.

In the analysis, the regions that exist in twenty states are included. Table no. A1 in the appendix gives details regarding the 103 regions, and the districts that form these regions. Figures A1 to A3 in the appendix give an idea of the homogeneity of the regions. Identifying whether convergence and

spatial convergence exist across these 'clusters', that is sub-state regions, becomes essential from the perspective of state policy planning.

2. Regional Income

The districts are the basic units used to create regions in this study. Therefore, to understand the regional income, the income at the district level should be aggregated at the regional level. Per-worker regional domestic product is used to identify the regional income. A per-worker regional domestic product also represents the productivity of the region.

Indicus Analytics provides information about the domestic product at a district level for the years 2001-2015. The aggregated Gross Regional Domestic Product (GRDP) is derived from aggregating the Gross District Domestic Product (GDDP) (Current Prices). The worker population at the district level is determined from the division of GDDP by GDDP per worker. This worker population at the district level is aggregated to compute the worker population at the regional level. A per-worker Gross Regional Domestic Product (GRDP) at current prices is derived from the aggregated regional population and GRDP. The equations (1) to (3) explain the process of calculating the per worker GRDP.

Equation 1

$$\mathrm{GRDP_A} = \sum_{i=1}^{n} \mathrm{GDDP_i} \tag{1}$$

Where,

$\mathrm{GDDP_i}$ is a Gross District Domestic Product of 'i' district in region 'A',

"n" is a total number of the districts in region "A",

$\mathrm{GRDP_A}$ is a Gross Regional Domestic Product of region "A"

Equation 2

$$\mathrm{Worker_A} = \sum_{i=1}^{n} \frac{\mathrm{GDDP_i}}{\text{Per Worker } \mathrm{GDDP_i}} \tag{2}$$

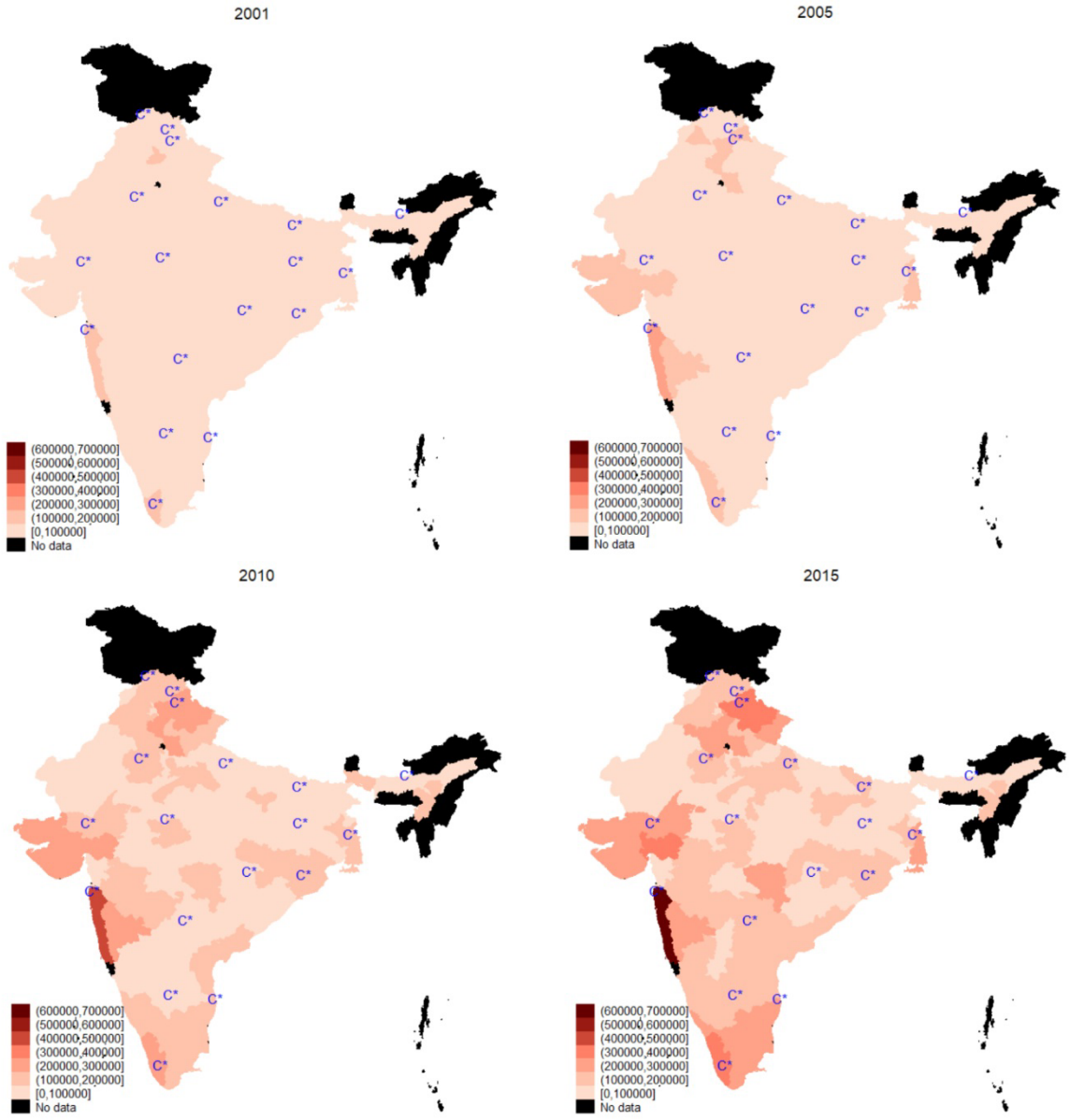
Where, $\mathrm{GDDP_i}$ is a Gross District Domestic Product of 'i' district in region 'A',

"n" is a total number of the district in region 'A',

$\mathrm{Worker_A}$ is a total number of workers in region 'A'

Equation 3: $Per\ Worker\ GRDP_A = \frac{GRDP_A}{Worker_A}$ (3)

Figure 1: Per Worker Regional Domestic Product



An interactive map depicts how per worker GRDP changes over the period of time across the regions (Figure 1).

- Regions with a state capital in a given state begin to improve in terms of per worker GRDP.
- In the latter stages, the neighbouring region also shows improvement.
- Gujarat and Chhattisgarh are the exceptions; in these states, the initial growth in per-worker GRDP occurs in regions other than the state capital. This is because both these regions have some cities that are historically involved in industrial (Bilaspur in the Bilaspur region of Chhattisgarh) and trade-related activities (Porbandar, Jamnagar, Bhavnagar, Ahmedabad in the Saurashtra region of Gujarat).

- As is seen in regions with a state capital, the regions neighbouring Bilaspur and Saurashtra show an improvement in per-worker GRDP.
- The remote regions that don't have a state capital are either picking up late or still not picking up in terms of per-worker GRDP.

Clearly, there is a spillover of wealth from one region to another. Figure 1 shows that while early increases in per-worker GRDP may be seen in specific regions (such as those with the state capital or those with a significant history of economic activity), the neighbouring regions are also benefiting from this growth. There's a lot to learn from this spatial pattern. Therefore, it is important to comprehend how regional income and neighbouring regions' income are moving and whether they are converging or not. This study thus includes regions within a state to highlight how convergence works across sub-state regions within a state. It also implements spatial panel data analysis to incorporate the spatial aspects of convergence.

3. Convergence

Empirical identification of convergence can be done by using the β-regression model (Durlauf and Quah, 1999). It gives the estimation for Beta convergence, which refers to a growth phenomenon where poor regions grow faster than rich regions (Sala-i-Martin, 1996). The empirical findings of growth models support the theoretical argument that convergence is possible (Maddison, 1991 and Barro-Lee, 2001 data set). Region-specific studies and cross-country studies do identify the presence of convergence. Table 1 describes cross-country studies and region-specific studies for Neoclassical and Augmented Neoclassical theories.

Table 1: Empirical Studies for Neoclassical and Augmented Neoclassical theories

Empirical studies	Country/Region	Approach	Result
Barro, et al., (1991) Barro and Sala-i-Martin (1992)	United States of America	Augmented neoclassical growth model where human capital is included in the model.	β-convergence across states which suggests poorer regions within country tends to grow faster than richer regions
Barro and Sala-i-Martin (1992)	Japanese Prefectures and United States of America	Augmented neoclassical growth model where human capital is included in the model.	β-convergence across states and prefectures.
Cashin (1995)	Australian Colonies	Neoclassical growth model	Divergence across colonies tends to decline.
Sala-i-Martin (1996)	European Countries OCED Countries	Neoclassical growth model	Convergence with different speed for different periods. Divergence for few periods.
Armstrong (1995)	European Countries	Neoclassical growth model	Convergence with declining rate.
Persson (1995)	Sweden	Neoclassical growth model	Convergence.
Cashin and Sahay (1996)	India States	Neoclassical growth model	Convergence.
Singh, et al (2010)	Indian districts	Neoclassical growth model	Convergence to steady state

While empirical studies based on Neoclassical and Augmented Neoclassical growth theories show convergence, empirical studies based on theories critical to Neoclassical growth theories (including disequilibrium theories and New Endogenous Growth theories) show the opposite. Table 2 summarises the empirical studies.

Table 2: Empirical Studies for Disequilibrium theories as well New Endogenous Growth theories

Empirical studies	Country/Region	Approach	Result
Quah (1996a) Quah (1996b)	European Countries United States	Critical to empirical approach of Neoclassical Growth Model due to non-inclusion of spill-over effects.	Ambiguous Result.
Marjit and Mitra (1996)	Indian States	Critical to empirical approach of Neoclassical Growth Model	Divergence.
Rao, Shand and Kalirajan (1999)	India States	Modification in Augmented Neoclassical growth model by adding population related variables.	Divergence.
Sachs, et al (2002)	India States	Incorporating agricultural reforms in growth equation	Divergence.
Rey and Montouri (1999)	United States	Inclusion of spill-over effects.	Convergence due to spill-over effect
Arbia and Piras (2005)	European Regions	Inclusion of spill-effect	Convergence
Sardadvar (2012)	European Regions	Inclusion of spatial dependence	Convergence
Shaban (2006)	Maharashtra Districts	Spatial Convergence	Regional convergence.
Kocornik-Mina (2009)	Indian States	Spatial Convergence	Divergence
Kalra and Thakur (2015)	Indian States	Spatial Convergence	Divergence

Above empirical works talk about how the spill-over effect affects the growth and income of a region. These works include the spill-over effect from regions and convergence across regions. The research question of how spill-over effects can be included in the framework of convergence is not investigated explicitly in the Indian regional context. This research work tries to fill the gap by adapting spatial convergence analysis for Indian sub-state regions. In this section, in addition to the traditional approach of convergence by β-regression, spatial convergence is also investigated. Spatial convergence is a process where the differences in income across regions adjacent to each other in space

tend to decrease. The panel data model incorporates the spatial aspects as well as the fixed effect to control the time-invariant characteristics of the regions.

3.1 Panel data beta convergence

Mathematically, the growth equation for convergence for panel data can be written as[3]:

Equation 4*:*

$$ln\left[\frac{y_{i,t+k}}{y_{i,t}}\right] = \alpha + \beta \times ln\, y_{i,t} + \sum_{i=2}^{N} \delta_i \times R_i + \varepsilon_{i,t} \quad (4)$$

Where, i (i=1,2,3,4,...,N) represents regions and t (t=1,2,3,4,..,T) represents the time periods. ln $\ln\left[\frac{y_{i,t+k}}{y_{i,t}}\right]$ is the annual growth rate of per worker gross regional domestic product (per worker GRDP) of region 'i' for a time period t-k. $\ln y_{i,t}$ can be interpreted as initial per worker GRDP for the given time period. R_i is a dummy variable for region 'i' and δ_i is its coefficient. R_i is included to incorporate the region-specific effects. If β is positive, then divergence is happening across regional growth rate and if β is negative then beta convergence that is conditional convergence is happening across regional growth rate

3.2 Panel data spatial beta convergence

As **Error! Reference source not found.** shows the spatial pattern, the spatial aspect should also be investigated. The equation 4 can be modified to incorporate the spatial aspect.

Equation 5

$$\ln\left[\frac{y_{i,t+k}}{y_{i,t}}\right] = \alpha + \beta \times \ln y_{i,t} + \sum_{i=2}^{N} \delta_i \times R_i + \rho \sum_{j=1}^{N} w_{ij} \times \ln y_{j,t} + \varepsilon_{i,t} \quad (5)$$

w_{ij} is an element from the binary spatial weights matrix (W), which is one if region i and region j are the neighbouring regions. A binary spatial weights matrix W represents the relationship between neighbour regions i and j. It has zero in diagonal. The neighbouring regions of region i are defined as j regions that have w_{ij} =1. w_{ij} is one only if regions are sharing borders otherwise it is zero. $\varepsilon_{i,t}$ is independently and identically distributed. It is also assumed that it has no spatial autocorrelation.

The model specification can be done to incorporate spatial autocorrelation.

Equation 6

$$\ln\left[\frac{y_{i,t+k}}{y_{i,t}}\right] = \alpha + \beta \times \ln y_{i,t} + \sum_{i=2}^{N} \delta_i \times R_i + \rho \sum_{j=1}^{N} w_{ij} \times \ln y_{j,t} + \lambda \sum_{j=1}^{N} w_{ij} \times \varepsilon_{i,t} + u_{i,t} \quad (6)$$

Equation 6 incorporates the spatial autocorrelation. The appropriate model is selected by adopting the approach given by LeSage and Pace (2009), Belotti, et al. (2013), and Elhorst (2014). Spxtregress stata package is used to estimate the spatial models (Kapoor, et al., 2007; StataCorp, 2017). Following

Elhorst (2014), diagnostic tests are also done to select the appropriate model (Appendix 7.2 discusses this in detail).

As a fixed effect is involved, adding the dummy for state capital regions will be omitted due to its time-invariant nature. Therefore, to check the convergence across non-state capital regions separately, separate models are used for 'all regions' and 'regions without state capital'. The spatial model for regions with state capital can't be implemented as the weight matrix is in binary form, and most of the regions with state capitals don't share boundaries with each other. ρ shows the impact of initial neighbouring regions' income on a given region's growth for a given time period. If it is negative, then it shows negative spill-over, which means rich income regions have negative impact on the neighbouring region's growth; if it is positive, rich income regions have positive impact on the neighbouring region's growth (positive spill-over).

Table 3: Conditions for β-convergence and spatial convergence

ρ>0	β>0	Spatial convergence with beta divergence with
ρ>0	β<0	Spatial convergence with beta convergence
ρ<0	β>0	Spatial divergence with beta divergence
ρ<0	β<0	Spatial divergence with beta convergence

4. Result

Table 4: Descriptive Statistics

Variable	Observation	Mean	Std. Dev.	Min	Max
Per worker GRDP	1442	81743.02	66748.04	588.01	570814.9
Log (Per worker GRDP)	1442	10.93	1.06	6.38	13.255
Annual Growth rate of per worker GRDP	1442	0. 079	0 .098	-0.116	2.053
Regions	1442	52	29.742	1	103
Year	1442	2008.5	4.033	2001	2014

Table 1 shows the descriptive statistics. The annual growth rate for the region has a negative lower-bound value.

Figure 2: Histogram with Normal Density Plot

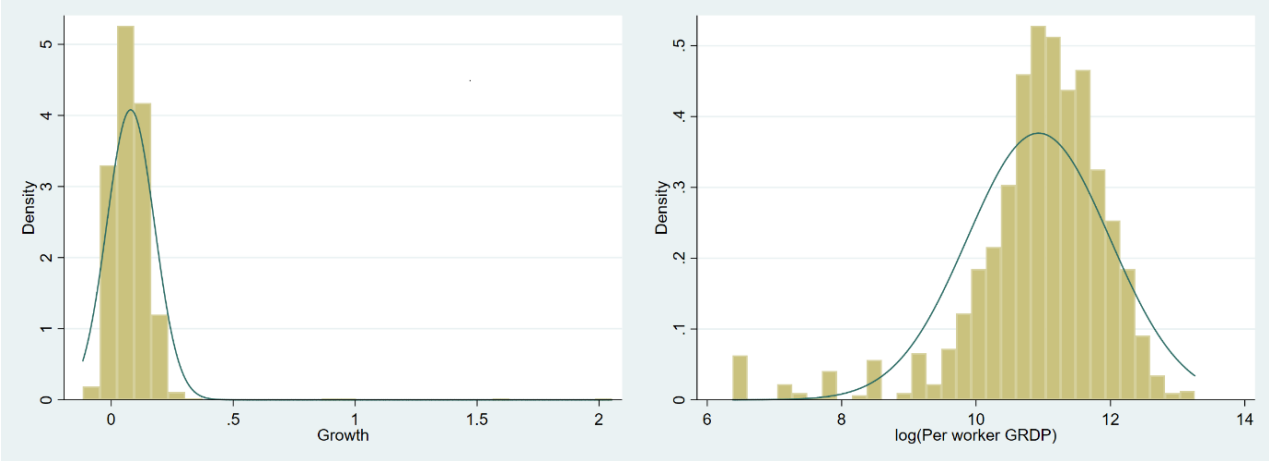


Figure 2 shows histogram which does suggest that panel data growth rate[4] and log (Per Capita GRDP) tend to be normally distributed.

Table 5: Panel data Fixed Effect Model

	All regions (FE model)	Regions with state capital (FE model)	Regions without state capital (FE model)
log (Per-Worker GRDP)	-0.06***	-0.08***	-0.06***
	(0.006)	(0.02)	(0.004)
Constant	0.79***	0.94***	0.74***
	(0.06)	(0.24)	(0.05)
R2 within	0.0846	0.0527	0.1460
R2 between	0.0406	0.5846	0.3260
R2 overall	0.0038	0.0384	0.0019
F value	123.68	12.23	190.96
(P value)	(0.00)	(0.00)	(0.00)
Number of Observations	1442	238	1204
Number of Groups	103	17	86
Hausman test:			
chi2	123.02	6.39	244.14
(P value)	(0.00)	(0.01)	(0.00)
Test for region specific			
effect:	2.26	1.84	4.47
F value	(0.00)	(0.03)	(0.00)
(P value)			

Table 2 shows panel data regression for convergence. A fixed effect model is an appropriate model over random effect and pooled models. The coefficients of log(Per-Worker GRDP) in all three models are statistically significant and negative, which suggests convergence. The coefficient of log (Per-Worker GRDP) for regions with a state capital is more negative than the coefficient of log (Per Worker GRDP) for overall regions as well as other regions, which suggests the convergence of growth rate across regions with state capital is faster compared to convergence across all regions.

Table 6: Spatial Panel Model

	All regions (Spatial Durbin Error FE model)	Regions without state capital (Spatial Durbin Error FE model)
log (Per-Worker GRDP)	-0.11***	-0.08***
	(0.011)	(0.008)
Spatial lag of log (Per-Worker GRDP)	0.01***	0.005**
	(0.002)	(0.002)
λ	0.06***	0.12***
	(0.007)	(0.007)
Log-likelihood	1314.4530	1689.3619
Number of Observations	1442	1204
Number of Groups	103	86
AIC value	-2620.906	-3370.724
Hausman test:		
chi2	90.80	117.62
(P value)	(0.00)	(0.00)
Wald test of spatial terms:		
chi2	88.36	268.39
(P value)	(0.00)	(0.00)

The Spatial Durbin Error Fixed Effect model is appropriate over the Spatial Lag Fixed Effect model[5]. In the spatial model, the coefficient of log(per-worker GRDP) is negative, which does suggest the convergence across region. Adding to that, the coefficient of spatial lag of log(per-worker GRDP) is positive, which means the positive spillover of per worker regional GRDP. This suggests that the region with a high per-worker regional GRDP leaves a positive impact on the neighbouring regions' growth.

5. Concluding remarks

In this empirical study, the primary focus was placed on analysing the convergence of gross regional domestic product (GRDP) per worker across 103 Indian regions over a course of time spanning from 2001 to 2015. The findings of a study using fixed-effect panel data on 103 regions, including 17 regions with state capitals, provide evidence of convergence.

The process of convergence is stronger for the regions with state capitals compared to regions without state capitals. The spatial analysis has also provided some insight regarding the convergences. It is found that there is a positive spill-over impact of GRDP per worker. The "rich regions" are able to boost the economic growth of their neighbouring regions. The interactive map also points out the spill-over of per worker GRDP.

The spatial analysis indicates that there are two types of convergence: β-convergence across all regions, and β-convergence among neighbouring regions. This study provides clear evidence for spatial convergence in the Indian context. As spatial dependence with a positive spill-over effect of per-worker GRDP is observed, it will be interesting to understand the phenomena through which this spill-over effect is happening.

References

Alesina, A., Bagir, R., & Easterly, W. (1999). Public goods and ethnic divisions. The Quarterly journal of economics, 114(4), 1243-1284.

Arbia, G., & Piras, G. (2005). Convergence in per-capita GDP across European regions using panel data models extended to spatial autocorrelation effects.

Arbia, G., Basile, R. and Piras G. (2005), "Using Spatial Panel Data in Modelling Regional Growth and Convergence", ISAE Working Paper No. 55, Rome: ISAE.

Armstrong, H. W. (1995). Convergence among regions of the European Union, 1950–1990. Papers in Regional Science, 74(2), 143-152.

Barro, R. J., & Lee, J. W. (2001). Barro-Lee Education Data.

Barro, R. J., Sala-i-Martin, X., Blanchard, O. J., & Hall, R. E. (1991). Convergence across states and regions. Brookings papers on economic activity, 107-182.

Barro, R. T., & Sala-i-Martin, X. (1992). Regional growth and migration: A Japan-United States comparison. Journal of the japanese and International Economies, 6(4), 312-346.

Barro, R. J., & Sala-i-Martin, X. (1995). Economic Growth McGraw-Hill. New York.

Belotti, F., Hughes, G., & Mortari, A. P. (2013). XSMLE: Stata module for spatial panel data models estimation. Statistical software components S, 457610.

Cashin, P. (1995). Economic growth and convergence across the seven colonies of Australasia: 1861–1991. Economic Record, 71(2), 132-144.

Cashin, P., & Sahay, R. (1996). Regional economic growth and convergence in India. Finance & Development, 33(001).

Durlauf, S. N., & Quah, D. T. (1999). The new empirics of economic growth. Handbook of macroeconomics, 1, 235-308.

Elhorst, J. P. (2014). Spatial econometrics: from cross-sectional data to spatial panels (Vol. 479, p. 480). Heidelberg: Springer.

Hirschman, A. O. (1958). The strategy of economic development. New Haven: Yale University Press.

Kalra, R., & Thakur, S. (2015). Development patterns in India: Spatial convergence or divergence?. GeoJournal, 80(1), 15-31.

Kapoor, M., Kelejian, H. H., & Prucha, I. R. (2007). Panel data models with spatially correlated error components. Journal of econometrics, 140(1), 97-130.

Kocornik-Mina, A. (2009). Spatial econometrics of multiregional growth: The case of India. Papers in Regional Science, 88(2), 279-300.

LeSage, J., & Pace, R. K. (2009). Introduction to spatial econometrics. Chapman and Hall/CRC.

Maddison, A. (1991). Dynamic forces in capitalist development: A long-run comparative view. Oxford University Press, USA.

Malgavkar, P. D., & Ghiara, B. M. (1969). Regional development: Where and how. In L. k. Sen (Ed.), Readings of micro level planning and rural growth centers. Hyderabad: National Institute of Community Development.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. The quarterly journal of economics, 107(2), 407-437.

Marjit, S., & Mitra, S. (1996). Convergence in regional growth rates: Indian research agenda. Economic and Political Weekly, 2239-2242.

Persson, J. (1995). Convergence in per capita income and migration across the Swedish counties 1906-1990. IIES.

Quah, D. T. (1996a). Regional convergence clusters across Europe. European economic review, 40(3-5), 951-958.

Quah, D. T. (1996b). Empirics for economic growth and convergence. European economic review, 40(6), 1353-1375.

Rao, M. G., Shand, R. T., & Kalirajan, K. P. (1999). Convergence of incomes across Indian states: A divergent view. Economic and Political Weekly, 769-778.

Rey, S. J., & Montouri, B. D. (1999). US regional income convergence: a spatial econometric perspective. Regional studies, 33(2), 143-156.

Sachs, J. D., Bajpai, N., & Ramiah, A. (2002). Understanding regional economic growth in India. Asian Economic Papers, 1(3), 32-62.

Sala-i-Martin, X. X. (1996). The classical approach to convergence analysis. The economic journal, 1019-1036.

Sardadvar, S. (2012). Growth and disparities in Europe: Insights from a spatial growth model. Papers in Regional Science, 91(2), 257-274.

Schaeffer, M. (2013). Can competing diversity indices inform us about why ethnic diversity erodes social cohesion? a test of five diversity indices in Germany. Social Science Research, 42(3), 755-774.

Shaban, A. (2006). Regional structures, growth and convergence of income in Maharashtra. Economic and Political Weekly, 1803-18.

Singh, N., Kendall, J., Jain, R. K., & Chander, J. (2010). Regional Inequality in India in the 1990s: Trends and Policy Implications, Development Research Group Study No. 35, Department of Economic Analysis and Policy, Reserve Bank of India, Mumbai.

Solow, R. M. (1956). A contribution to the theory of economic growth. The quarterly journal of economics, 70(1), 65-94.

StataCorp, L. L. C. (2017). Stata spatial autoregressive models reference manual.

Swan, T. W. (1956). Economic growth and capital accumulation. Economic record, 32(2), 334-361.

Appendix

A1. Homogeneity

The degree of diversity that exists in a society can be measured by its ethnic fractionalization (Alesina et al., 1999). Schaeffer (2013) also use the ethnic diversity index, which is computed by deducting the conventional Hirschman-Herfindahl Index (HHI) (Hirschman, 1958) from one. This gives the ethnic diversity index. In order to have a better understanding of the presence of homogeneity, the HHI index is used. Based on the categories of caste, religion, and caste-religion that are available in the NFHS-4 data, three different HHI concentration indices are constructed. Values that are higher indicate a higher concentration, which may be interpreted as a sign that the regions are homogeneous.

$$\text{HHI (Caste Based)}_i = \sum_{i=1}^{4} x_i^2 \tag{A1}$$

Where, x_i is a share of caste "i" and caste categories are:
Scheduled Caste, Scheduled Tribes, Other Backward Communities and Don't know.

$$\text{HHI (Religion based)}_i = \sum_{i=1}^{9} y_i^2 \tag{A2}$$

Where, y_i is a share of religion "i" and religion categories are:
Buddhism, Christian, Hindu, Jain, Muslim, Parsi, Sikh, Other Religion, No Religion.

$$\text{HHI (Caste Based)}_i = \sum_{i=1}^{4} z_i^2 \tag{A3}$$

Where, z_i is a share of categories formed by religion and caste categories which are 36 in total.

The caste-based concentration is seen in the figure A1. The caste-based concentration for the states is shown in the left panel, while the caste-based concentration for the regions defined in the research is displayed in the right panel. The concentration level is not dropping for regions compared to states where few regions have higher concentration.

Figure A 1: Caste Based Concentration

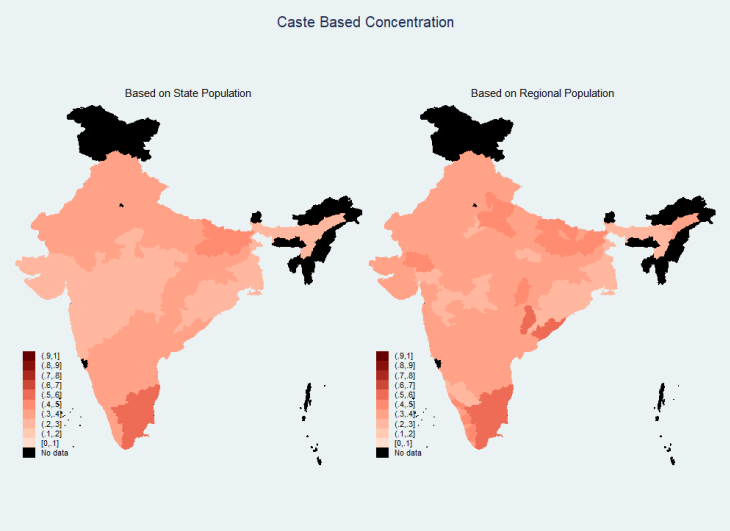


Figure A2 shows the religion-based concentration. Similar to caste-based concentration, map is getting "redder" for regions compared to states which show few regions have higher concentration.

Figure A 2: Religion Based Concentration

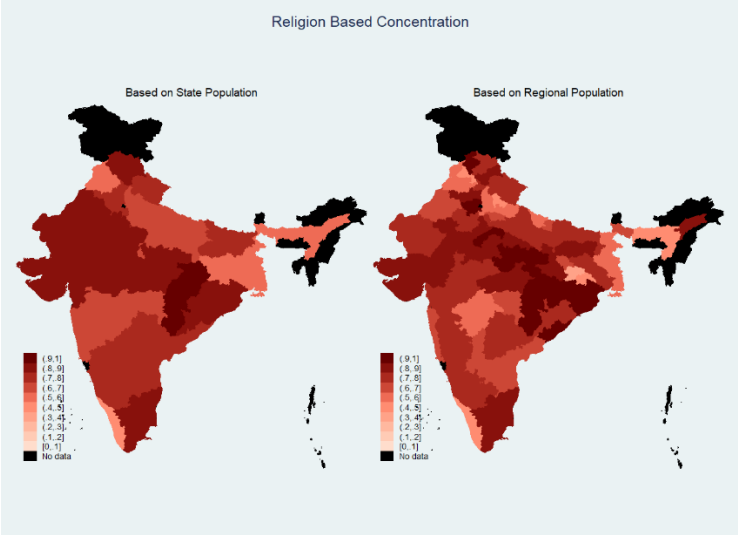
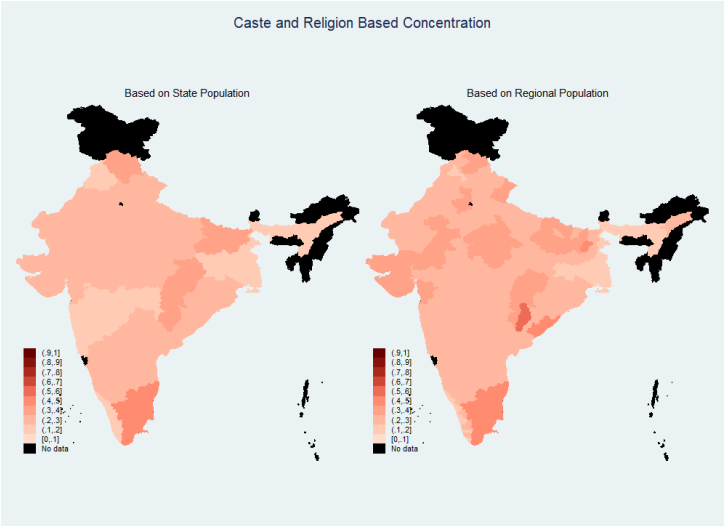


Figure A3 shows the religion-caste-based concentration. Due to more categories, the indices value naturally will be lower. But pattern is still similar to figure A1 and figure A2 Similar to caste-based concentration, map is getting "redder" for regions compared to states which show few regions have higher concentration.

Figure A 3: Caste and Religion Based Concentration



Following the concentration maps, the Indian regions formed in the given study do have homogeneity. These regions do possess both the important characteristics: administrative characteristics and homogeneity.

Table A 1: Regions

State	Region	Districts	Reason behind region-formation
Andhra Pradesh	Coastal Andhra Region	East Godavari Guntur Krishna Prakasam S.P.S. Nellore West Godavari	Physiographical regions, Political regions, Historical division
	Rayalaseema Region	Anantapur Chittoor Kadapa YSR Kurnool	Physiographical regions, Political regions, Historical division
	Uttarandhra Region	Srikakulam Visakhapatnam Vizianagaram	Physiographical regions, Political regions, Historical division
Assam	Hills and Barak Valley	Cachar Hailakandi Karbi Anglong Karimganj North Cachar Hil	Political regions, Historical division
	Lower Assam	Baksa	Political regions,

		Barpeta	Historical division
		Bongaigaon	
		Chirang	
		Dhubri	
		Goalpara	
		Kamrup	
		Kamrup (Metro)	
		Kokrajhar	
		Nalbari	
	North Assam	Darrang	Political regions,
		Marigaon	Historical division
		Nagaon	
		Sonitpur	
		Udalguri	
	Upper Assam	Dhemaji	Political regions,
		Dibrugarh	Historical division
		Golaghat	
		Jorhat	
		Lakhimpur	
		Sibsagar	
		Tinsukia	
Bihar	Bhagalpur	Banka	Political regions,
		Bhagalpur	Administration division
	Darbhanga	Darbhanga	Political regions,
		Madhubani	Administration division
		Samastipur	
	Kosi	Madhepura	Political regions,
		Saharsa	Administration division
		Supaul	
	Magadh	Arwal	Political regions,
		Aurangabad	Administration division
		Gaya	
		Jehanabad	
		Nawada	
	Munger	Begusarai	Political regions,
		Jamui	Administration division
		Khagaria	

		Lakhisarai Mungair Sheikapura	
	Patna	Bhabhua / Kaimur Bhojpur Buxar Nalanda Patna Rohtas	Political regions, Administration division
	Purnea	Araria Katihar Kishanganj Purnea	Political regions, Administration division
	Saran	Gopalganj Saran Siwan	Political regions, Administration division
	Tirhut	Champaran (East) Champaran (West) Muzaffarpur Sheohar Sitamarhi Vaishali	Political regions, Administration division
Chhattisgarh	Bastar	Bastar Dantewara	Political regions, Administration division
	Bilaspur	Bilaspur Janjgir Korba Raigarh	Political regions, Administration division
	Durg	Durg Kawardha Rajnandgaon	Political regions, Administration division
	Raipur	Bijapur Dhamtari Kanker Mahasmund Narayanpur Raipur	Political regions, Administration division

	Surguja	Jashpur Koriya Surguja	Political regions, Administration division
Gujarat	Central Gujarat	Ahmedabad Anand Dahod Kheda Panchmahal Vadodara	Physiographical regions, Political regions, Historical division
	North Gujarat	Banaskantha Gandhinagar Mehsana Patan Sabarkantha	Physiographical regions, Political regions, Historical division
	Saurashtra - Kutch	Amreli Bhavnagar Jamnagar Junagadh Kutch Porbandar Rajkot Surendranagar	Physiographical regions, Political regions, Historical division
	South Gujarat	Bharuch Dangs Narmada Navsari Surat Tapi Valsad	Physiographical regions, Political regions, Historical division
Haryana	Ambala	Ambala Kurukshetra Panchkula Yamunanagar	Political regions, Administration division
	Faridabad	Faridabad	Political regions,

		Mewat Palwal	Administration division
	Gurugram	Gurgaon Mahendragarh Rewari	Political regions, Administration division
	Hisar	Fatehabad Hissar Jind Sirsa	Political regions, Administration division
	Karnal	Kaithal Karnal Panipat	Political regions, Administration division
	Rohtak	Bhiwani Jhajjar Rohtak Sonepat	Political regions, Administration division
Himachal Pradesh	Kangra	Chamba Kangra Una	Political regions, Administration division
	Mandi	Bilashpur Hamirpur Kullu Lahul & Spiti Mandi	Political regions, Administration division
	Shimla	Kinnaur Shimla Sirmaur Solan	Political regions, Administration division
Jharkhand	Kolhan	Sariakela / Kharsawan Singhbhum East Singhbhum West	Political regions, Administration division
	North Chotanagpur	Bokaro Chatra	Political regions, Administration division

		Dhanbad Giridih Hazaribagh Khodrama / Koderma Ramgadh	
	Palamu	Gadva / Garhwa Latehar Palamau	Political regions, Administration division
	Santhal Pargana	Devghar / Deogarh Godda Jamtara Pakund / Pakur Sahebganj Santhal Paragana / Dumka	Political regions, Administration division
	South Chotanagpur	Gumla Khunti Lohardagga Ranchi Simdega	Political regions, Administration division
Karnataka	Belagavi	Bagalkote Belgaum Bijapur Dharwad Gadag Haveri Uttara Kannada	Political regions, Administration division
	Bengaluru	Bangalore (Rural) Bangalore (Urban) Chikkaballapur Chitradurga Davanagere	Political regions, Administration division

		Kolar Ramanagaram Shimoga Tumkur	
	Gulbarga	Bellary Bidar Gulbarga Koppal Raichur Yadagiri	Political regions, Administration division
	Mysuru	Chamaraja Nagar Chickmagalur Dakshina Kannada Hassan Kodagu Mandya Mysore Udupi	Political regions, Administration division
Kerala	Central Kerala	Eranakulam Malappuram Palakkad Thrissur	Political regions, Administration division
	North Kerala	Kannur Kasaragod Kozhikode Wayanad	Political regions, Administration division
	South Kerala	Alappuzha	Political regions, Administration division

		Pathanamthitta Idukki Kollam Kottayam Thiruvananthapuram	
Madhya Pradesh	Bhopal	Bhopal Raisen Rajgarh Sehore Vidisha	Political regions, Administration division
	Chambal	Bhind Morena Sheopur Kalan	Political regions, Administration division
	Gwalior	Ashoknagar Datia Guna Gwalior Shivpuri	Political regions, Administration division
	Indore	Alirajpur Barwani Burhanpur Dhar Indore Jhabua Khandwa Khargone	Political regions, Administration division

Jabalpur	Balaghat Chhindwara Dindori Jabalpur Katni Mandla Narsinghpur Seoni	Political regions, Administration division
Narmadapuram	Betul Harda Hoshangabad	Political regions, Administration division
Rewa	Rewa Satna Sidhi Singrauli	Political regions, Administration division
Sagar	Chhatarpur Damoh Panna Sagar Tikamgarh	Political regions, Administration division
Shahdol	Anuppur Shahdol Umaria	Political regions, Administration division
Ujjain	Dewas	Political regions, Administration division

		Mandsaur Neemuch Ratlam Shajapur Ujjain	
Maharashtra	Khandesh	Ahmednagar Dhule Jalgaon Nandurbar Nasik	Political regions, Administration division
	Konkan	Mumbai sub Mumbai sub Raigad Ratnagiri Sindhudurg Thane	Political regions, Administration division
	Marathwada	Aurangabad Beed Hingoli Jalna Latur Nanded Osmanabad Parbhani	Political regions, Administration division
	Paschim Maharashtra	Kolhapur	Political regions, Administration division

		Pune Sangli Satara Solapur	
	Vidarbha (Nagpur)	Bhandara Chandrapur Gadchiroli Gondia Nagpur Wardha	Political regions, Administration division
	Vidarbha (Varhad)	Akola Amarawati Buldhana Washim Yeotmal	Political regions, Administration division
Odisha	Central Revenue Division (Cuttack)	Balasore Bhadrak Cuttack Jagatsinghapur Jajapur Kendrapara Khurda Mayurbhanja Nayagarh Puri	Political regions, Administration division, Revenue

	Northern Revenue Division (Sambalpur)	Angul Bargarh Bolangir Deogarh Dhenkanal Jharsuguda Keonjhar Sambalpur Sonepur Sundargarh	Political regions, Administration division, Revenue
	Southern Revenue Division (Berhampur)	Boudh Gajapati Ganjam Kalahandi Koraput Malkangiri Nawarangpur Nuapada Phulbani (Kandhamal) Rayagada	Political regions, Administration division, Revenue
Punjab	Doaba	Hoshiarpur Jalandhar	Political regions, Administration division

		Kapurthala S.B.S Nagar	
	Majha	Amritsar Gurdaspur Taran Taran	Political regions, Administration division
	Malwa	Barnala Bhatinda Faridkot Ferozpur Ludhiana Mansa Moga Patiala Sangrur Shri Mukatsar Sahib	Political regions, Administration division
	Poadh	Fatehgarh Sahib Roopnagar S.A.S Nagar	Political regions, Administration division
Rajasthan	Ajmer	Ajmer Bhilwara Nagaur Tonk	Political regions, Administration division
	Bharatpur	Bharatpur Dholpur Karoli Swami Madhopur	Political regions, Administration division

Bikaner	Bikaner Churu Ganganagar Hanumangarh	Political regions, Administration division
Jaipur	Alwar Dausa Jaipur Jhunjhunu Sikar	Political regions, Administration division
Jodhpur	Barmer Jaisalmer Jalore Jodhpur Pali Sirohi	Political regions, Administration division
Kota	Baran Bundi Jhalawar Kota	Political regions, Administration division
Udaipur	Banswara Chittorgarh Dungarpur Pratapgarh Rajsamand Udaipur	Political regions, Administration division

Tamil Nadu	Tamil Nadu	Ariyalur Chengalpattu MGR / Kancheepuram Chennai Chidambanar / Toothukudi Coimbatore Dharmapuri Dindigul Anna Kanyakumari Karur Krishnagiri Madurai Nagapattinam Namakkal North Arcot / Vellore Perambular Periyar (Erode) Pudukkottai Ramananthapuram Salem Sivagangai / Pasumpon South Arcot / Cuddalore Thanjavur The Nilgiris Theni Thirunelveli Thiruppur Thiruvallur Thiruvannamalai Tiruchirapalli / Trichy Tiruvarur Villupuram Virudhunagar / Kamarajar	Political regions
Telangana	Telangana	Adilabad Hyderabad	Political regions

		Karimnagar Khammam Mahabubnagar Medak Nalgonda Nizamabad Rangareddy Warangal	
Uttar Pradesh	Agra division	Agra Firozabad Mainpuri Mathura	Political regions, Administration division
	Aligarh division	Aligarh Etah Hathras Kasganj/Khansi Ram Nagar	Political regions, Administration division
	Ayodhya division	Ambedkar Nagar Barabanki Faizabad Sultanpur	Political regions, Administration division
	Azamgarh division	Azamgarh Ballia Mau	Political regions, Administration division
	Bareilly division	Bareilly Budaun Pilibhit Shahjahanpur	Political regions, Administration division

Basti division	Basti Santh Kabir Nagar Sidharthnagar	Political regions, Administration division
Chitrakoot division	Banda Chitrakoot Hamirpur Mahoba	Political regions, Administration division
Devipatan division	Bahraich Balrampur Gonda Shravasti	Political regions, Administration division
Gorakhpur division	Deoria Gorakhpur Kushi Nagar / Padrauna Mahrajgani	Political regions, Administration division
Jhansi division	Jalaun Jhansi Lalitpur	Political regions, Administration division
Kanpur division	Auraiya Etawah Farrukhabad Kannauj Kanpur Dehat	Political regions, Administration division

	Kanpur Nagar	
Lucknow division	Hardoi Kheri Lucknow Rae-Bareily Sitapur Unnao	Political regions, Administration division
Meerut division	Bagpat Buland Shahar G.B.Nagar Ghaziabad Meerut	Political regions, Administration division
Mirzapur division	Mirzpur Santh Ravi Das Nagar / Bhadoi Sonbhadra	Political regions, Administration division
Moradabad division	Amroha/J.B.Fulenagar Bijnor Moradabad Rampur	Political regions, Administration division
Prayagraj division	Allahabad Fatehpur Kushambi Pratapgarh	Political regions, Administration division

	Saharanpur division	Muzaffarnagar Saharanpur	Political regions, Administration division
	Varanasi division	Chandauli Ghazipur Jaunpur Varanasi	Political regions, Administration division
Uttarakhand	Garhwal	Chamoli Dehradun Garhwal Haridwar Rudraprayag Tehri Garhwal Uttar Kashi	Political regions, Administration division
	Kumaon	Almorah Bageshwar Champavat Nainital Pithorgarh Udham Singh Nagar	Political regions, Administration division
West Bengal	Burdwan division	Birbhum Burdwan Hooghly	Political regions, Administration division
	Jalpaiguri division	Cooch Behar Darjeeling Jalpaiguri	Political regions, Administration division

Malda division	Dakshin Dinajpur Malda Murshidabad Uttar Dinajpur	Political regions, Administration division
Medinipur division	Bankura East Midnapore / Purba Midnapore Jhargram Purulia West Midnapore	Political regions, Administration division
Presidency division	24-Paraganas North 24-Paraganas South Kolkata Nadia	Political regions, Administration division

A.2: Spatial Models and diagnostic test

Spatial Model can be represented as

$$Y = \rho WY + \beta X + WX\theta + \varepsilon$$

where β is parameters for exogenous explanatory variables which denotes the K×1 vector, ρWY represents the endogenous interaction effect and $WX\theta$ represents the exogenous interaction effects. Y is vector of dimension N×1, X represents an N×K matrix and $\varepsilon = \lambda W\varepsilon + u$

If $\rho = 0$ then,

$$Y = \beta X + WX\theta + \varepsilon$$

Above equation represents Spatially Lagged explanatory variables (SLX) where there is no endogenous interaction but there is a presence of exogenous interaction.

If $\theta = 0$ and where $\lambda \neq 0$ then Spatial Error Model (SEM) and if λ is also equal to 0 then OLS: $Y = \beta X + \varepsilon$

In equation, $Y = \rho WY + \beta X + WX\theta + \lambda W\varepsilon + u$

If $\rho = 0$, $\theta \neq 0$ and $\lambda \neq 0$ then model is known as Spatial Durbin Error Model (SDEM),

If $\rho = 0$, $\theta \neq 0$ and $\lambda = 0$ then model is known as Spatial Lagged Model (SLX),

If $\rho = 0$, $\theta = 0$ and $\lambda \neq 0$ then model is known as Spatial Error Model (SEM),
If $\rho = 0$, $\theta = 0$ and $\lambda = 0$ then model is known as OLS,
If $\rho \neq 0$, $\theta = 0$ and $\lambda = 0$ then model is known as Spatial Autoregressive Model (SAR),
If $\rho \neq 0$, $\theta = 0$ and $\lambda \neq 0$ then model is known as Spatial Autoregressive Model with Autoregressive Disturbances (SARAR),
If $\rho \neq 0$, $\theta \neq 0$ and $\lambda = 0$ then model is known as Spatial Durbin Model (SDM),
If $\rho \neq 0$, $\theta \neq 0$ and $\lambda \neq 0$ then model is known as General Nesting Spatial Model (GES),
Based on the condition of ρ, λ and θ, the appropriate model can be selected among above specifications (LeSage and Pace, 2009; Elhorst, 2014).

NOTES

[1]Refer to Sala-i-Martin (1996).
[2]Poor regions grow faster than rich regions.
[3]Refer to Arbia et al (2005).
[4]There are outliers in growth rate. But these outliers are non-influencing as dropping them is not causing any significant changes in the result.
[5]Following Elhorst (2014), appropriate model is selected.