Growth Divergence between Indian States

., Kaustubh and Ghosh, Taniya

Reserve Bank of India, Indira Gandhi Institute of Development Research

1 January 2024

Growth Divergence between Indian States[1]

Dr. Taniya Ghosh[2], Kaustubh[3]

Abstract

The paper investigates the causes of rising inequality among Indian states in terms of per capita State Domestic Product (GDP) in recent decades by examining the convergence/divergence between 20 major Indian states from 2000 to 2019. The paper adds to the existing literature by including the newly created states of Jharkhand, Chhattisgarh, and Uttarakhand in the sample. The paper, like previous research, finds 'conditional' convergence between states per capita GDP; however, the literature cannot determine whether this is due solely to different steady states caused by state-specific characteristics or to differences in productivity growth rates. According to our findings, differences in productivity growth rates, as well as different steady states, are the drivers of states' increasing per capita GDP inequality. Factors such as population growth rate, bank credit/GDP ratio, and shares of agriculture and industry in the GDP explain the differences in steady states between states, which drive the disparity in per capita GDP between states. Therefore, to reduce inequality, the states with low per capita GDP may benefit from rebalancing these variables. Furthermore, we must also investigate the causes of variations in productivity growth rates to address growing regional inequality.

Keywords – *Solow Model, Conditional Convergence, Beta-convergence*, *Stochastic convergence, Unit Root, Fixed Effect,*

---

[1] The views expressed in the paper are of the authors and they not necessarily reflect the views of RBI or IGIDR
[2] Corresponding Author, Indira Gandhi Institute of Development Research (IGIDR), Gen. A. K. Vaidya Marg, Film City Road Mumbai, 400065, India, Email Add.: taniya@igidr.ac.in, Tel.:91-22-28426536, ORCID ID: https://orcid.org/0000-0002-9792-0967
[3] Department of Economic and Policy Research, Reserve Bank of India, Mumbai, Email- kaustubh@rbi.org.in

1. Introduction

Exploring the reasons for the deepening of regional inequality is recent decades is a key policy question for equitable growth of the Indian economy. Several works in recent years have explored the reasons behind the growing regional income inequality in India. Ghosh and Kaustubh (2023), Misra et al. (2022), and Misra (2019) have focused on finding the source of growing inequality by looking at the divergence in the factors of production (i.e. capital accumulation or total factor productivity growth). However, another possible direction of exploration in this regard, is testing the convergence hypothesis. This paper focuses on testing the convergence hypothesis between Indian states.

The convergence hypothesis is derived from the Solow model, which predicts that poorer economies will grow faster than richer economies. This is because, they have lesser per capita capital, initially, and hence have a higher return on capital due to the model's assumption of diminishing marginal returns on capital. Hence, they will eventually catch up with the richer economies by achieving higher per capita income. Thus, according to the convergence hypothesis, poorer states must grow faster than richer states.

However, in the scatter plot below (Figure 1), there is a positive relationship between the variables rather than the negative relationship predicted by the convergence hypothesis, with the log of per capita GDP[4] of the states in 2000[5] on the x-axis and growth in per capita GDP of states between 2000 and 2019 on the y-axis. Hence, rather than convergence, per capita GDPs are showing divergence. To explain such divergence, the Solow model uses the conditional convergence hypothesis, which eliminates the assumption that the states have identical socio-

[4] In the study GDP of states denotes their Gross State Domestic Product.
[5] For this analysis, a year, say 2000, includes represents the financial year 2000-01

economic characteristics. Therefore, if the conditional convergence hypothesis is true, states will show convergence conditioned on various socio-economic variables.

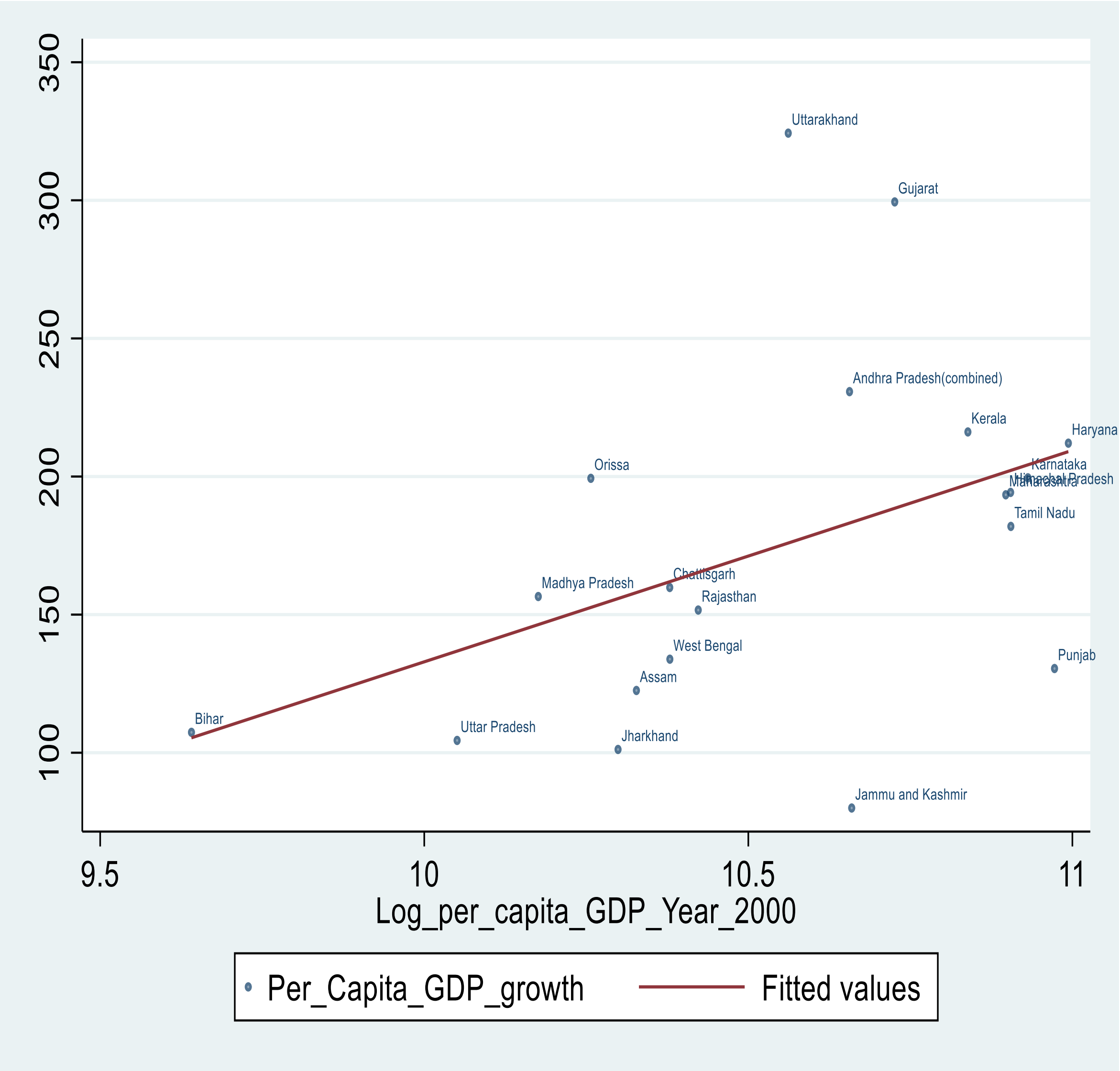


Figure 1: The relationship between per capita GDP growth and log of initial per capita output

By examining the existence of unconditional convergence and conditional convergence of per capita GDP between 20 major Indian states in the time-period 2000-2019[6] and exploring the

[6] To avoid the impact of outlier events like COVID on results, the time-period of the study is limited to 2019-20.

reasons driving the divergence in per capita GDP between the Indian states, the study investigates reasons driving the increasing inequality in per capita GDP of Indian states.

The literature largely agrees on the existence of conditional convergence between Indian states but is unclear whether this conditional convergence is caused solely by different steady states or also by differences in productivity growth rates. This clarification has important policy implications. If the divergence is solely due to different steady states caused by the level effect variables such as investment rate, elasticity of capital, population growth, etc, the inequality between states can be reduced by rebalancing the states' relative positions in these variables. However, if the disparity is also due to differences in productivity growth rates, we must investigate the reasons for these variations in productivity growth rates. In this regard, conditional beta-convergence and stochastic convergence tests can throw some light.

In the unconditional beta-convergence, growth rates of the per capita GDP of states over the study period are regressed on previous or initial time-period per capita GDP of states (Ghosh et al. (1998) and Sanga et.al. (2017)). For the conditional beta-convergence, the regression is augmented with additional explanatory variables, such as socio-economic characteristics, state and time fixed effects, etc.

For stochastic convergence, the conditional convergence hypothesis is tested by checking the stationarity of the ‘relative’ GDP of the states. The relative GDP is the log ratio of a state’s per capita GDP to national per capita GDP (Misra et al. (2018) and Ghosh (2008)). The presence of unit roots in the relative GDP of states’, indicates that the effect of unequal productivity shocks on the relative GDP persists over time and that the state's relative GDP does not revert to its mean value in the future. Thus, when the outcomes of conditional beta-convergence and stochastic divergence are combined with the presumption that all states are close to their steady states, it can be concluded that the divergence is caused by variations in productivity shocks between the states rather than just differences in steady states (Bernard et al. (1996)).

Moreover, because panel estimations of conditional convergence include a lag term in explanatory variables such as previous period per capita GDP, there is a possibility of bias in the fixed effect estimation due to the correlation between error terms and regressors with the lag term (Islam (1995)). This potential bias, however, was not considered in some of the literature when estimating convergence using panel data techniques. Therefore, we use dynamic panel techniques to analyse the conditional convergence estimates.

Following the Solow model framework, based on finding of presence of conditional beta-convergence and stochastic divergence of per capita GDP of Indian states, our study indicates that differences in the productivity growth rates along with different steady states are the drivers of the states' increasing per capita GDP inequality. Particularly, the factors like the population growth rate, bank credit/GDP ratio, and shares of agriculture and industry in the GDP explain the different steady states between the states, and hence the divergence in per capita GDP between states. The low per capita GDP states should focus on rebalancing these variables to reduce inequality. Furthermore, diminishing returns associated with a higher credit/GDP ratio and a higher (lower) share of industry (agriculture) in GDP indicate that, compared to higher per capita GDP states, improving them will have a much greater impact on GDP growth in low per capita income states. In addition, the divergence between the states is a result of variations in productivity growth rates and, we must investigate the causes of these variations to reduce regional inequality.

The remainder of the study is organized as follows. In Section 2, we discuss the literature on testing the unconditional and conditional convergence hypothesis in the per capita GDP among the Indian states. In Section 3, we explain the literature gap and motivation. Research methodology is explained in Section 4. The empirical findings are discussed in Section 5, and the contribution of the study, concluding remarks and policy implications are presented in Section 6 and Section 7.

2. Literature Survey

There is a wealth of literature on testing the convergence hypothesis and investigating possible causes of divergence among Indian states. In terms of the methodologies used, literature can be broadly summarized into parametric methods and non-parametric methods. The parametric methods are regression based where we assume the residuals follow a normal distribution. The parametric approaches can be further classified into sigma, beta, and stochastic convergence testing.

In the sigma-convergence, a time-series plot of various measures of inequality of the per capita GDP of states such as the Gini coefficient, etc., was analysed. If the per capita GDP inequality measure was increasing over time, then it was sigma divergence and sigma-convergence if it was decreasing with time (Ghosh (2008) and Sanga et al. (2017)). The method had also been augmented by regressions of measures of inequality of state per capita GDP with a linear or quadratic time trend. Additionally, some studies added slope and intercept dummies indicating the time event in the regression to check impact of certain events, such as the 1991 reforms, on sigma-convergence (Ghosh (2008), Ghosh et al. (1998) and Sanga et al. (2017)).

In beta-convergence, we check if there is negative relationship between per capita GDP in the start of time-period and per capita GDP growth in entire time-period. In the unconditional beta-convergence test, the negative coefficient of previous or initial per capita GDP of states in the regression of growth rates of the per capita GDP indicates a beta-convergence. For conditional beta-convergence, the regression was augmented with additional explanatory variables, such as production structural variables, socio-economic characteristics, state, and time fixed effects, etc. Several techniques, including cross-sectional regressions, and panel regressions such as fixed effect estimations, dynamic panel methods, instrument variables regressions, etc., were used to estimate the regressions. beta-convergence (Cherodian et.al. (2015), Nagraj et al. (1998), Lolayekar et al. (2020) and Ghosh (2008)).

For stochastic convergence, the conditional convergence hypothesis was tested by checking the stationarity of the 'relative' GDP of the states. Therefore, we would have unconditional convergence if the relative GDPs of states are stationary around 0, and convergence with different steady states if the relative GDP of states is stationary around values other than 0. If the relative GDP of states' is non-stationary, it means that the productivity shock to the states' per capita GDP persists, and the states' relative GDPs do not revert to their mean.

In terms of the results, most of the literature have indicated sigma-divergence and conditional beta-convergence of Indian states in terms of per capita GDP. The conditional beta-convergence was conditioned on various socio-economic characteristics of states, such as the share of industry and agriculture in GDP, credit growth, literacy, public expenditure as a share of GDP, etc. In addition, many papers indicated that after the 1991 reforms, the sigma divergence between states' per capita income increased.

The literature also indicated convergence between the states in social indicators such as IMR (Infant Mortality Rate), MMR (Maternal Mortality Rate), literacy rates, etc unlike the divergence seen in per capita GDP (Ghosh (2008)). In terms of sectors of the economy, there was convergence in per capita output in the agricultural sector but divergence in the per capita output of the industrial and manufacturing sectors among Indian states (Sanga et al. (2017)).

Although there is a large body of literature on testing convergence hypotheses between Indian states using a diverse set of methodologies. Some literature gaps are identified and transcribed in the next section.

3. Literature Gap and Motivation

The three states created in 2000, Jharkhand, Uttarakhand, and Chhattisgarh were not included separately in most studies' samples and were combined with their parent states, Bihar, Uttar Pradesh, and Madhya Pradesh. Additionally, they have begun the period of study before their creation. However, it is crucial to test the convergence hypothesis in the time-period following their creation in addition to separating them from their parent states in the sample of the states to determine whether there is divergence in the time-period following their creation.

The literature largely agrees on the existence of conditional convergence between states, but it is unclear whether this conditional convergence is caused solely by different steady states or also by variations in rates of productivity growth between the states. This clarification has significant policy implications. If the divergence between the states is only because of different steady states caused by the level effect variables such as population growth rate, investment rate, and depreciation rate, inequality between the states can be reduced by simply rebalancing the states' relative positions in these variables. However, if the divergence between the states is also a result of variations in productivity growth rates, we must investigate the causes of these variations to reduce regional inequality.

The conditional beta-convergence test can tell if there is conditional convergence or divergence in the per capita GDP but cannot specifically tell if the conditional convergence is based only on different steady states or is also due to different productivity growth rates. Though in the conditional convergence estimations, explanatory variables explaining productivity growth, such as social and physical infrastructure, are taken as control variables but the estimations do not explicitly test if the divergence is only due to different steady states or there is a difference in the productivity growth rates. On the other hand, in the stochastic convergence test, the presence of unit roots in the relative GDP of states will indicate that the effect of unequal productivity shocks on the relative GDP persists over time and that the relative GDP of the state does not revert to its mean steady-state value in the future.

Inferring that the divergence is caused by differences in persistent productivity growth shocks between the states rather than just differences in steady states is, therefore, possible when the results of conditional beta-convergence and stochastic divergence are combined with the assumption that all states are close to their steady states (Bernard et al. (1996)). Thus, we can use the combination of the conditional beta-convergence test and stochastic convergence test to indicate that the divergence is also due to the differences in productivity growth.

As panel estimations of the conditional convergence contain a lag term in the explanatory variables such as per capita GDP of the previous period, there is a possibility of bias in the fixed effect estimations due to the correlation between error term and regressors due to lag term (Islam (1995)). The impact of this potential bias, however, was not considered in some literature while estimating convergence using panel data techniques.

Therefore, it is advantageous that we also use dynamic panel techniques like System GMM and first difference GMM to confirm our estimates of conditional convergence. In the first difference GMM, estimation equation is changed to the first difference form in the first step of the estimation process to eliminate state fixed effects. Then, the lagged levels of the first differenced endogenous explanatory variables are taken as the instruments to correct the bias (Arellano et al. (1991)). In the system GMM, to correct the poor finite sample properties of the first difference GMM, the first difference equation is augmented by a level equation to get a system of equations. Then the lagged first differences of the explanatory variables act as instruments for the level equation and the lagged level explanatory variables act as instruments for the first difference equation (Arellano et al. (1995) and Blundell et al. (1995)).

The conditional convergence estimations in the literature do not consider several other possible explanatory variables, such as the export orientation of the state's economy and the amount of FDI it has received, etc., which can influence the divergence between the states' per capita GDP.

Also, estimations in the literature did not focus on exploring the possibility of diminishing returns to the impact of explanatory variables on per capita GDP in the conditional convergence estimation. For instance, the effect of improving the credit/GDP ratio on per capita GDP may have a diminishing return because at low levels of credit/GDP, improving the ratio can result in a greater increase in per capita GDP than at high levels of credit/GDP. By examining whether applying a log transformation to the explanatory variables enhances the model by increasing the R square, we can test for the diminishing return. The log-transformation provides a parsimonious way to account for the diminishing returns as other methods, like including quadratic and higher-order transformations of the explanatory variables in the regression, will lead to loss of degrees of freedom and multi-collinearity between explanatory variables.

As a result of the diminishing return of explanatory variable on per capita GDP growth, even modest improvements in the explanatory variables for low per capita GDP states relative to higher per capita GDP states will have a significant impact on reducing the inequality between the states.

4. Methodology

The methodology of the study can be briefly summarised in the following 3 steps.

4.1 Step 1

We will check the panel stationarity of log of relative GDP of the states by the Im-Pesaran-Shin Panel unit-root test at 0,1 and 2 lags with subtracting the cross-sectional mean from the variable. Im-Pesaran-Shin test is an extension of Augmented Dickey Fuller test of stationarity in panel time series. To control of cross-sectional dependence, we have run the test after subtracting the cross-sectional means from the variable. If the relative GDP is found to be stationary at 0 then, we will have evidence for unconditional convergence. If the relative GDP of the states is non-stationary, then we will need the methodologies in Step 2 to see if there is conditional convergence or divergence.

4.2 Step 2

The estimation of the convergence equation is done in the study using a slight modification of the Barro style convergence regression (Barro et al. (1992)). In the Barro style regression, the growth rate of the per capita GDP of the states in the given period is taken as the dependent variable and the initial per capita GDP of the states is the explanatory variable. The coefficient of the explanatory variable is used to calculate the speed of convergence in the sample. To account for heterogeneity between the states time-invariant dummies for states have been introduced in the regression. The Equation 1 below provides an equation-based explanation of the Barro regression.

$$\frac{1}{T}\ln\left[\frac{y_{i,t0+T}}{y_{i,t0}}\right] = \alpha_i - \left[\frac{1-e^{-\beta T}}{T}\right]\ln\left(y_{i,t0}\right) + u_{i,t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1)},$$

where y stands for real per capita GDP, β is the convergence coefficient, α_i is time-invariant but varies across i (states). Thus, α_i is an economy index which is a function of steady-state output per effective labour and technical progress, T is the length of the time-period, t is a time index, t_0 is the initial time and $u_{i,t}$ is an independently distributed error term. In the pooled model of regression estimation, we assume that all states have the same economy index α_i and same speed of convergence β. In the fixed-effect model with time dummies, we assume states have a different economic index α_i but the same speed of convergence β.

However, we have made some modifications to Equation 1 in the estimation of the convergence equation. We are estimating the AR (1) model of the log of GDP per capita in the panel data with an annual time period (like Hayashi (2000)). Thus, instead of taking the first difference of the log of GDP per capita as the dependent variable as in Equation 1, we have taken the log of GDP per capita as the dependent variable. It allows us to introduce the productivity shocks common to all states each year, by adding time dummies to the estimation equation. The Equation 2 describes the estimation equation that is estimated.

$lny_{i,t} = \phi * lny_{i,t-1} + \alpha_i + \beta_t + u_{i,t}$ ………………………………………………. (2).

In Equation 2, y is the per capita GDP, α_i is the time-invariant state fixed effect for state i, β_t is the time fixed effect denoting common productivity shock at time t, and u is the idiosyncratic error. Hence, if estimated ϕ > 1, we have beta divergence between the states, and if estimated ϕ < 1, we have conditional beta-convergence between the states.

To address the possibile bias in fixed effect estimation, we also estimated the equation using First Difference GMM and Systems GMM. In the GMM estimates, in addition to lags of the log of per capita GDP and lags of the first difference of log of per capita GDP, we will also include time dummies and state dummies as additional exogenous variables. Thus, in Step 2, we estimate the conditional convergence based on Barro-style regression in the following ways.

a. Pooled estimation to test for unconditional convergence,
b. Conditional convergence estimation to test for conditional convergence conditioned on time and state fixed effects,
c. Dynamic panel estimation using systems and difference GMM with dummies for state and time as additional instruments.

To account for heteroskedasticity, robust estimation is used for all the estimations.

4.3 Step 3

The causes of the divergence will then be explained by extracting time-invariant state fixed effects coefficients from the Step 2 estimations and then regressing state fixed effects on explanatory variables, like the share of industry and agriculture in the GDP, population growth, credit growth, export/GDP ratio, etc. After that, we will explore the possibility of diminishing returns of the explanatory variables on the per capita GDP by taking the log transform of the explanatory variables in the regression to explain the time-invariant state fixed effects better.

The data source in the study is the RBI handbook on Indian states for the years, 2021 and 2017.

5. Results

5.1 Step 1

The Figure 2 shows the panel data line plots of the states' relative GDP over the specified time-period. The graph demonstrates that the relative GDP for most of the states is non-stationary and is either trending upward or downward. Additionally, the plot demonstrates that the non-stationarity is not caused by any common structural break.

Table 1A in the Appendix displays the results of the Im-Pesaran-Shin Panel unit-root for panel stationarity of log of relative GDP of states. Presence of a unit root is the null hypothesis of the test, which is not being rejected at 0, 1 and 2 lags Thus, it supports our findings from Figure 2 about non-stationarity of the relative GDP.

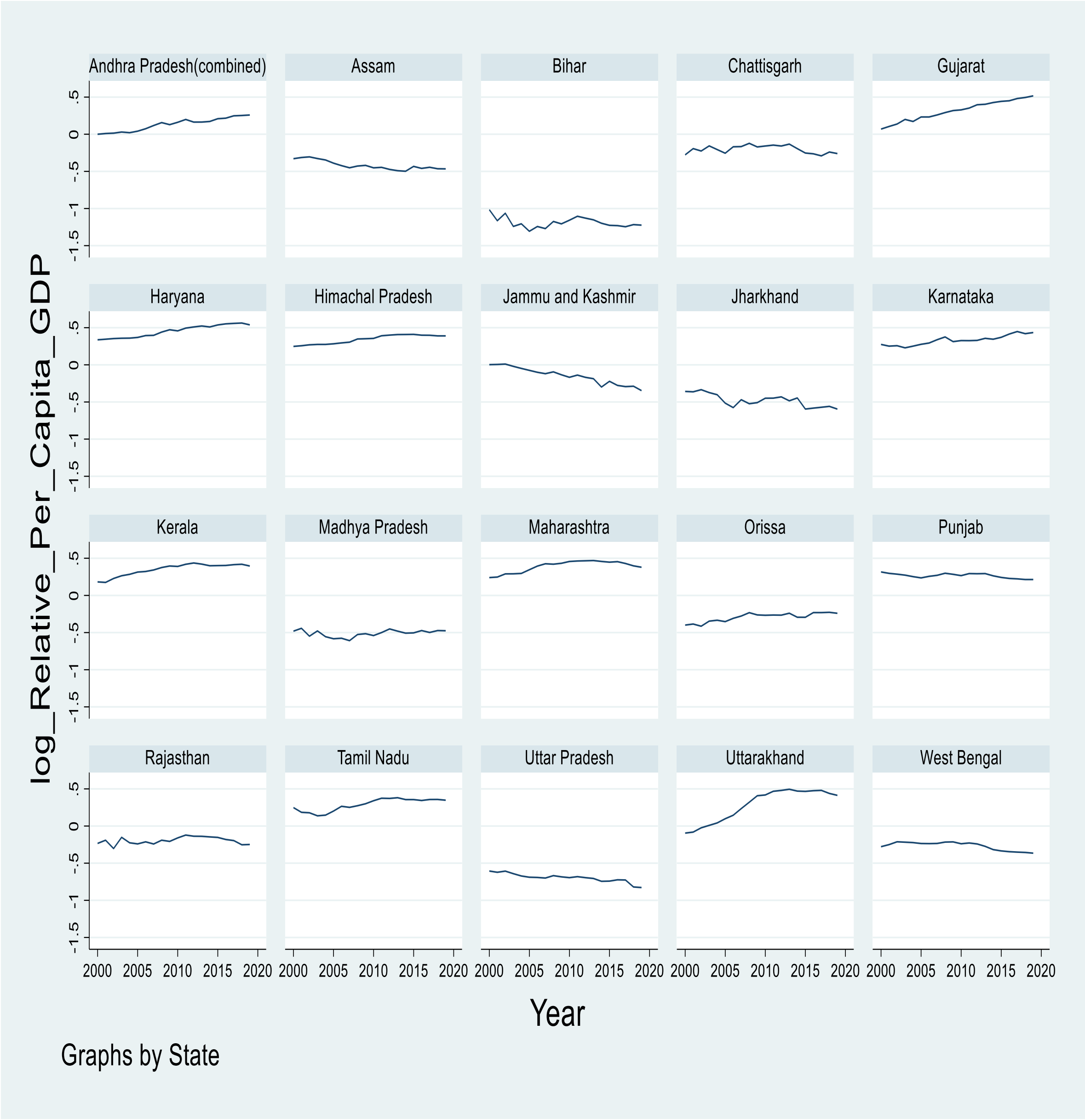


Figure 2: The panel data line plots of log of relative GDP of the states

<u>5.2 Step2</u>

The results of the conditional convergence estimation are summarised in the Table 1. After controlling for time and state fixed effects, the divergence result of the pooled model converts to convergence, as the value of the coefficient is 1.01 (greater than 1) in the pooled case but is 0.87 (less than 1) in the fixed effect case. The GMM estimates show that the bias in the fixed effect model was not large, and as the coefficient ϕ decreased only slightly, the speed of convergence increased slightly compared to panel estimation. This is because the sample's time-period (T = 20) is not short, as bias resulting from the presence of lagged explanatory variables in the panel estimation decreases as the time-period lengthens

Models	P-OLS in levels	FE	FD-GMM	S-GMM
Coefficient ϕ	1.01***	0.87***	0.81***	0.81***
R square	0.99	Within – 0.99 Between – 0.99 Overall – 0.99		

Table 1: Conditional convergence estimation

Note: ***** denotes 1% significance level, **denotes 5% significance level and *denotes 1% significance level**

Hence, the reason for the beta divergence of the states' per capita output growth rate are the time-invariant state fixed effect parameters. In addition, as discussed in the literature gap section, stochastic divergence, and conditional convergence under the Solow model framework, with the additional assumption that states are close to their steady states, further suggest that increasing inequality between the states is not just driven by their different steady states but also by the differences in their productivity growth.

5.3 Step 3

According to the Barro equation, the time-invariant state fixed effect, or in other words the time-invariant economic index of the state, is a function of the productivity growth and steady-state output per effective worker of the states. According to the Solow model, the steady-state output per effective worker of the states is a function of savings, elasticity of capital, population growth, depreciation rate, and productivity growth. Thus, to investigate the reasons for the divergence between the state's growth, we need to explore the causes behind the variations in the state fixed effect parameter.

To account for variations in state fixed effects, we have used three variables based on the Solow model. The structural index is the first. It is an index that is created using Principal Component Analysis (PCA) of the share of agriculture and industry in the state's GDP in 2000[7]. The two variables were chosen to create the index because a higher share of agriculture generally leads to low productivity growth as avenues of productivity growth are less in agriculture and a higher share of industries can lead to faster productivity growth (Nagraj (1998))[8]. Due to the strong negative correlation between the two variables, PCA was required. Therefore, separately including them in the regression would have resulted in high levels of multicollinearity among the explanatory variables. PCA is a statistical method to combine the variables into orthogonal linear combinations using the Eigen values and Eigen vectors of the correlation matrix of the variables. Then the calculated linear combinations containing the maximum variations of variables are taken as the indices representing those variables. (Chao et.al (2017)).

Thus, using PCA we have created an index named 'Structural Index' which is a linear combination of the share of agriculture and industry in GDP in the year 2000. It is the first Principal Component of the PCA which contains the largest variations of the variables (86 percent) among calculated linear combinations using the PCA. This is due a strong negative correlation between the two shares of the GDP. Additionally, the Eigenvectors in Table 3 show

[7] The ratio of Industry share to Agriculture share would have similar but simpler metric to depict the structural characteristic of the state's economy compared to taking PCA of Industry share and Agriculture share. Hence, for robustness test of model in the appendix we have run the regression similar to Model 3 (Table 3) with log of ratio of Industry share in GDP to Agriculture share in GDP as the explanatory variable instead log of structural index (Appendix Table 1B, Model A). The results for both regressions are similar, though log of structural index has slightly better explanatory power compared to log of ratio of industry share to agriculture share. Thus, we have continued with using structural index as the explanatory in the study.

[8] The reviewer has enquired about the potential role of share of services in GDP in explaining the divergence in per capita GDP between states. Regarding this, we can compare the results of regression with log of ratio of services share to agriculture share as the explanatory variable in addition to log of credit GDP ratio and population growth (Appendix Table 1B, Model B) and with regression in which we have taken the log of ratio of share of industry in GDP to the share of agriculture in GDP (Appendix Table 1B, Model A) as the explanatory variable instead of log of ratio of services share to agriculture share. The log of ratio of services share to agriculture share has little explanatory power as it's coefficient is insignificant with the P-value of 0.8 unlike the log of ratio of industry share to agriculture share whose coefficient is significant at 10 percent significance level.

that the weights of the two shares are nearly equal in the Structural Index in terms of magnitude. However, the weight of agriculture share is positive, and the weight of the industry share is negative. We have taken agriculture share and industry share of the year 2000 to create the structural index to make it exogenous to the growth of output per capita in the period 2000 to 2019.

Variable	Component 1	Component 2
Agriculture Share	0.71	0.71
Industry Share	-0.71	0.71

Table 2: Principal components (Eigenvectors)

The second variable is the Decadal Population growth between 2001 and 2011 (n) using 2011 and 2001 census data. The third variable is the ratio of credit given by scheduled commercial banks in the year 2001 and state GDP for 2001 is taken as a proxy for the private investment rate in the states. The credit data of the new states of Uttarakhand, Chhattisgarh, and Jharkhand is available on the RBI website from the year 2001. The data of the year 2001 is taken to calculate the credit/GDP ratio to make it as exogenous as possible to the growth rate of output per capita for the time-period 2000-2019.

5.3.1 Model 1: Simple linear model

The results of model 1 with time-invariant state fixed effect as the dependent variable and population growth rate (n), credit GDP ratio, and structural index as explanatory variables estimation have been summarised in Table 3. The signs of coefficients are according to the expectation. A higher structural index with positive weightage for agriculture share and negative weightage for industry share has a negative coefficient, as a higher share of agriculture and a lower share of industry is hypothesised to have lower productivity growth. Similarly, as predicted by the Solow model, higher population growth leads to lower growth, while higher credit to GDP ratio leads to higher growth. In terms of the magnitude of the coefficients the population growth rate seems to be more dominant in explaining the variations in the state fixed effect compared to the other two explanatory variables. However, the coefficients of the variables structural index and credit/GDP ratio are not significant at a 5% significance level. To better explain the variations in the state fixed effects, we will therefore introduce the possibility of diminishing returns in the relationship between the explanatory variables and the dependent variable.

5.3.2 Model 2: Log transformation of explanatory variables

To get a better understanding of the relationship between explanatory variables and the dependent variable we will investigate the scatter plots.

a. Between population growth (n) and state fixed effects

The scatter plot in Figure 1A in the Appendix indicates a negative linear relationship between state fixed effects and n. Thus, we do not need to make a non-linear transformation of the n in the regression model.

b. Between credit/GDP ratio and state fixed effect

The scatter plot in Figure 1B in the Appendix indicates a positive but non-linear relationship between state fixed effect and credit/GDP ratio. Hence, we have taken the log of credit/GDP ratio to account for the non-linearity and plotted the relationship between state fixed effects and credit/GDP ratio in Figure 1C in the Appendix. As we can see from Figure 1C after taking a log of the credit/GDP ratio, the relationship becomes linear. Thus, the credit/GDP ratio has a diminishing return on growth. Hence, we have taken log of the credit/GDP ratio as the explanatory variable in our regression instead of the credit/GDP ratio to improve the explanatory power of our model.

c. <u>Between state fixed effects and structural index</u>

The scatter plot in Figure 1D in the Appendix shows a possibility of a negative non-linear relationship between structural index and state fixed effect. However, because the index was constructed using PCA, it has negative values. Hence, we need to transform the index to an all-positive index to take the log of the index to model the non-linearity in the relationship between time-invariant state fixed effects and structural index. Hence, we have transformed the structural index to a range of values between 0 and 1 using max-min transformation as shown in equation 3.

$$TI = \frac{(\max\limits_{SI} - SI)}{(\max\limits_{SI} - \min\limits_{SI})} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3).$$

TI is Transformed Index, SI is Structural Index, $\max\limits_{SI}$ is the maxima of the Structural Index and $\min\limits_{SI}$ is minima of the Structural Index

The transformed index has a negative relationship with the structural index. This was done to model the transformed index to have a positive relationship with state fixed effects. Then we have taken log of the transformed index and plotted it in the scatter plot in Figure 1E in the Appendix with state fixed effects. Thus, improving the structure of the economy also has a diminishing return on growth. Hence, we have taken the log of the structural index as the explanatory variable in our regression instead of the structural index to improve the explanatory ability of our model. The results of regression estimation after taking the log of credit/ GDP ratio and log of the structural index as the explanatory variables are summarised in Table 3.

In comparison to Model 2, the log transformation of the variables significantly improved the R-square and the P values of the coefficients. We can therefore conclude that the log transformation of credit-GDP ratio and structural index explains the time-invariant state fixed effects better than the earlier regression without the log transformation. This confirms a diminishing return associated with the structural index or credit/GDP ratio of the state on GDP growth. Also, the signs of the coefficients are as expected. The coefficient of population growth rate (n) is negative while coefficients of log of credit/GDP ratio and log of the structural index are positive. In terms of the magnitude of coefficients, the population growth rate remains dominant in explaining the variations in the state fixed effect; however, its magnitude has decreased as compared to Model 1. Also, the log transformation of the structural index and credit GDP ratio has changed the interpretation of the coefficients of both variables in the regression. Now, their coefficients denote the change in state fixed effects due to a one percent change in both explanatory variables.

5.3.3 Model 3: Including export/GDP ratio as an explanatory variable

Exports have driven productivity growth in the Indian economy since the 1991 reforms due to increased global competition (Agarwal (2015)). Hence, we have added the export/GDP ratio of the states as an explanatory variable to check if a higher export orientation of the state's GDP leads to higher growth. The export/GDP percentage is taken from the NITI Aayog report on the Export Preparedness Index 2020. The data of this variable is difficult to find for the year 2000 as data for earlier years are difficult to find. Thus, there is a possibility of endogeneity bias in the regression estimation with export/GDP as an explanatory variable. The scatter plot in Figure 1F in the Appendix indicates a positive non-linear relationship between state fixed effects and export/GDP ratio.

Hence, we have taken the log of export/GDP ratio to account for the non-linearity in the relationship and plotted the relationship between state fixed effects and export/GDP ratio in Figure 1G in the Appendix. After taking the log of export/GDP ratio, we can see from Figure 1G that the relationship becomes linear. Thus, the export/GDP ratio seems to have a diminishing return on the growth. Hence, we have taken the log of the export/GDP ratio as the explanatory variable in our regression instead of the export/GDP ratio. The result of the regression is summarised in Table 3.

The positive sign of the coefficient of the log of export/GDP indicates a positive relationship between the export/GDP ratio and state fixed effects (which is also seen in Figures 1F and 1G in the Appendix). However, controlling for population growth rate (n), log of credit/GDP ratio, and log of the structural index, the log of export/GDP ratio does not have a significant impact on the time-invariant state fixed effects. Also, the inclusion of the export/GDP ratio as an additional explanatory variable in Model 3 does not change the R-square much compared to Model 2.

Model 1	
Variables	Coefficient
Constant	0.04
Population Growth Rate (n)	-0.61*
Credit/GDP Ratio	0.15**
Structural Index	-0.01
R-square- 0.54	
Model 2	
Variables	Coefficient
Constant	0.13***
N	-0.43**
Log Credit/GDP Ratio	0.04**
Log Structural Index	0.02***
R-square- 0.60	
Model 3	

Variables	Coefficient
Constant	0.13***
N	-0.38*
Log Credit/GDP Ratio	0.04**
Log Structural Index	0.02***
Log Export/GDP Ratio	0.003
R-square- 0.61	

Table 3: Regression Results of Model 1, Model 2, and Model 3

Note: *** denotes 1% significance level, **denotes 5% significance level and *denotes 10%significance level

6. Contributions of the study

We get a conditional convergence between the states as the introduction of the state fixed effects and time fixed effects lead to convergence in the per capita GDP growth. As the state fixed effects are time invariant, there is less possibility of endogeneity in the conditional convergence estimation where they are included in the estimation as the control variables. In addition, the use of dynamic panel techniques ensures that the bias due to the introduction of lagged dependent variable has been taken care of in establishing conditional convergence.

The stationary tests of the states' relative GDP show that under the Solow Model framework and the assumption that states are close to their steady states, the conditional convergence could be due to both different steady states and different productivity growth rates between states.

The explanatory variables based on the Solow model, such as the population growth, the share of agriculture and industry in the GDP, and the credit/GDP ratio can explain a large part of the variations in the time-invariant state fixed effects that are responsible for the divergence in growth rates between the states. For example, in terms of per capita GDP growth in the given time-period, the explanatory variable structural index can explain poor performance of agriculturally dependent states such as Bihar and Punjab and the good performance of industrialised states such as Gujarat. Similarly, the study explains why states with high rates of population growth, such as Bihar, Uttar Pradesh, etc., perform poorly in terms of per capita GDP growth.

The study demonstrated that the structural index and credit/GDP ratio has a diminishing return on growth. This indicates that there is a higher gain in terms of a declining share of agriculture and an increasing share of industry in the GDP at low levels of industry share and high levels of agriculture share, but the gain decreases as we move along the path. The impact of improving the credit/GDP ratio on growth is also much greater at low levels of credit/GDP ratio than it is at high levels of credit/GDP ratio.

The study also demonstrates that even though the time-invariant states fixed effects in the scatter plot have a positive relationship with the log of export/GDP ratio. The log of export/GDP ratio, however, has no significant effect on time-invariant states' fixed effects after controlling for the log of the structural index, population growth rate, and log of credit/GDP ratio.

7. Conclusion

Thus, we can conclude that though there is a conditional beta-convergence between the Indian states, the growing divergence between their per capita GDP will continue to grow if certain policy actions are not taken in the low per capita GDP states. Therefore, to catch up to the states with higher per capita GDP levels, they can work to increase the share of industry and decrease the share of agriculture in their GDP, improve the credit/GDP ratio and focus on decreasing the population growth rate.

However, the silver lining is that improvement in structural factors such as the share of agriculture and share of industry and credit/GDP ratio have a diminishing impact on growth. Therefore, improving them even slightly will have a much greater impact on their growth in low per capita income states. This can aid in tackling growing divergence between the Indian states. Furthermore, we must investigate the causes of variations in productivity growth rates across different states of India to reduce regional inequality. This can be a topic for further research.

References

- Agrawal, P. (2015). The Role of Exports in India's Economic Growth. *The Journal of International Trade & Economic Development*, 24(6), 835-859.

- Arellano, M., & Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error-components Models. *Journal of Econometrics*, 68(1), 29-51.

- Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2), 277-297.

- Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2), 223-251.

- Barro, R. J. & Sala-i-Martin, X. (1995). Economic Growth. McGraw-Hill, 1995, 539.

- Bernard, A. B., & Durlauf, S. N. (1996). Interpreting Tests of the Convergence Hypothesis. *Journal of Econometrics*, 71(1-2), 161-173.

- Blundell, R., & Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 87(1), 115-143.
- Cherodian, R., & Thirlwall, A. P. (2015). Regional Disparities in Per Capita Income in India: Convergence or Divergence? *Journal of Post Keynesian Economics*, 37(3), 384-407.
- Chao, Y. S., & Wu, C. J. (2017). Principal Component-based Weighted Indices and a Framework to Evaluate Indices: Results from the Medical Expenditure Panel Survey 1996 to 2011. *PloS One*, 12(9), e0183997.
- Ghosh, B., Marjit, S., & Neogi, C. (1998). Economic Growth and Regional Divergence in India, 1960 to 1995. *Economic and Political Weekly*, 1623-1630.
- Ghosh, M. (2008). Economic Reforms, Growth and Regional Divergence in India. *Margin: The Journal of Applied Economic Research*, 2(3), 265-285.
- Ghosh, M. (2012). Regional Economic Growth and Inequality in India during the Pre- and Post-reform Periods. *Oxford Development Studies*, 40(2), 190-212.
- Ghosh, T & Kaustubh (2023) Growth Decomposition of the Indian states using panel data techniques, *Applied Economics*. DOI: 10.1080/00036846.2023.2212974
- Islam, N. (1995). Growth Empirics: A Panel Data Approach. *The Quarterly Journal of Economics*, 110(4), 1127-1170.
- Hayashi F. (2000), Econometrics, *Princeton University Press*, pp. 342-346,
- Kar, S., Jha, D., & Kateja, A. (2011). Club-convergence and Polarization of States: A Nonparametric Analysis of Post-reform India. *Indian Growth and Development Review*, Emerald Group Publishing, vol. 4(1), pages 53-72, April.

- Lessmann, C., & Seidel, A. (2017). Regional Inequality, Convergence, and its Determinants–A View from Outer Space. *European Economic Review*, 92, 110-132

- Lolayekar, A. P., & Mukhopadhyay, P. (2020). Understanding Growth Convergence in India (1981–2010): Looking beyond the Usual Suspects. *PloS one*, 15(6), e0233549.

- Mishra, A., & Mishra, V. (2018). Re-examination of Convergence Hypothesis among Indian States in Panel Stationarity Testing Framework with Structural Breaks. *Applied Economics*, 50(3), 268-286.

- Misra, B.S. (2020), "Determinants of Total Factor Productivity in Indian States", *Indian Growth and Development Review*, Vol. 13 No. 1, pp. 259-282

- Misra, B. S., Nazlioglu, S., & Kucukkaplan, I. (2022). "Sources of Divergence in Income in Indian States", 2001–2015. *Development Policy Review*, 40, e12594. https://doi.org/10.1111/dpr.12594

- Nagaraj, R., Varoudakis, A., & Véganzonès, M. A. (2000). Long-run Growth Trends and Convergence across Indian States. *Journal of International Development: The Journal of the Development Studies Association*, 12(1), 45-70.

- NITI Aayog (2020), Export Preparedness Index (EPI) 2020

- Quah, D. (1992). Empirical Cross-section Dynamics in Economic Growth. *European Economic Review*, Volume 37, Issues 2–3, 1993, Pages 426-434,

- Quah, D. (1993). Galton's Fallacy and Tests of the Convergence Hypothesis. *The Scandinavian Journal of Economics*, 427-443.

- Sanga, P., & Shaban, A. (2017). Regional Divergence and Inequalities in India. *Economic & Political Weekly*, 52(1), 103.

Appendix

S.No.	Lags	w_t_bar-statistics	P-values
1.	0	-0.05	0.48
2.	1	0.07	0.52
3.	2	-0.48	0.32
Table 1A: Results of Panel Stationarity Test of Relative GDP			

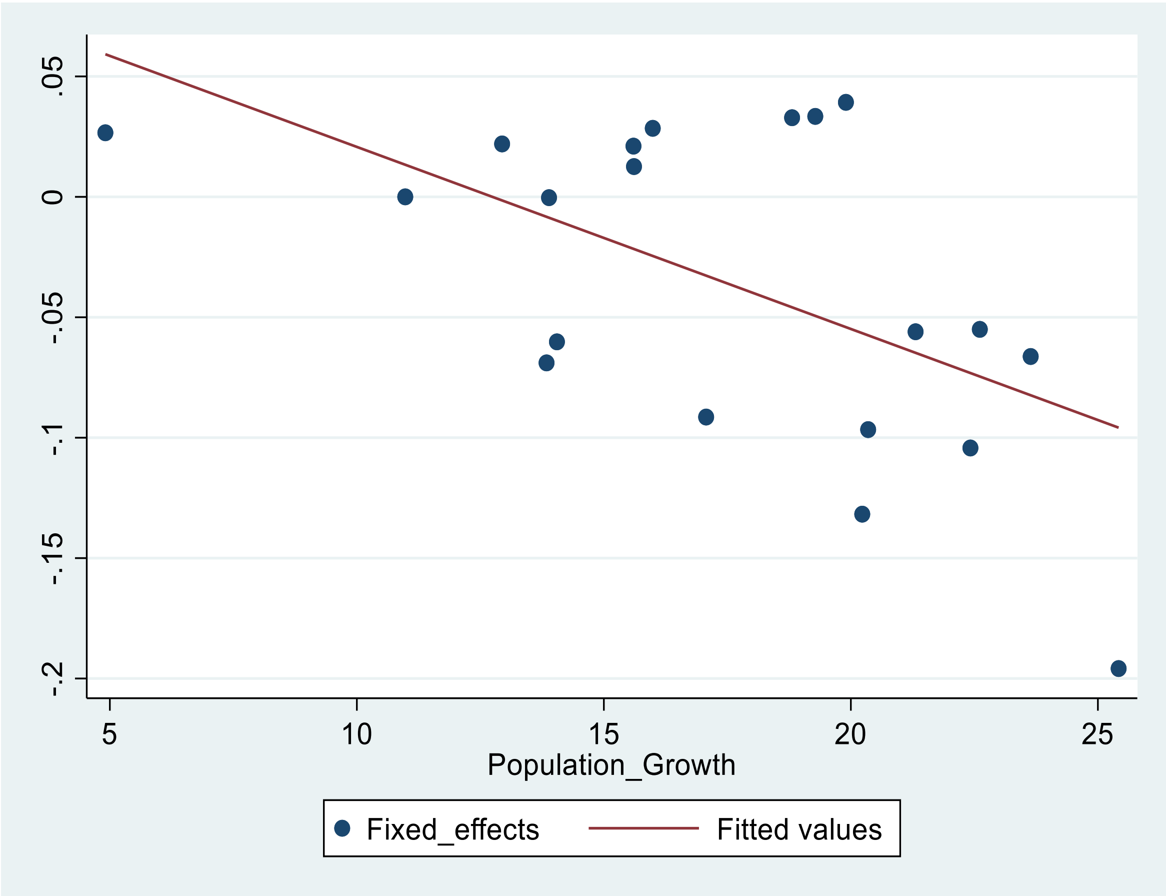


Figure 1A: The scatter plot indicates a negative linear relationship between State Fixed Effects and Population Growth Rate

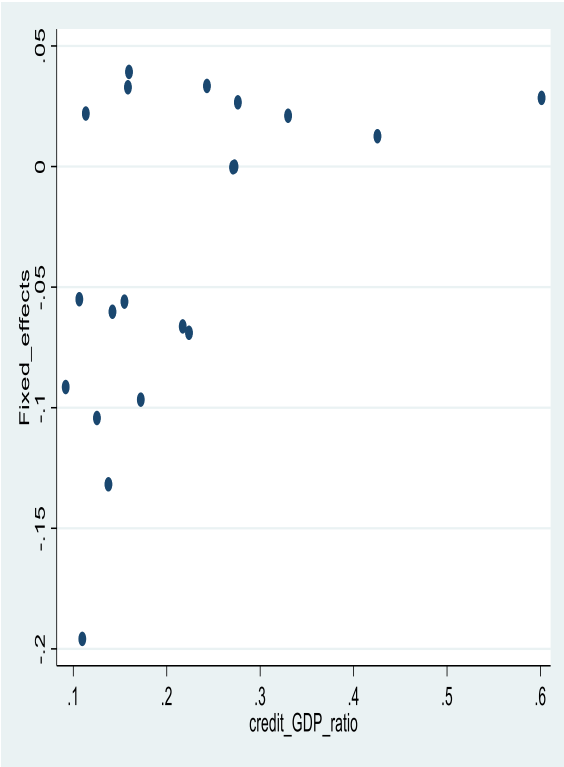


Figure 1B: Scatter plot between States Fixed Effect and Credit/GDP ratio

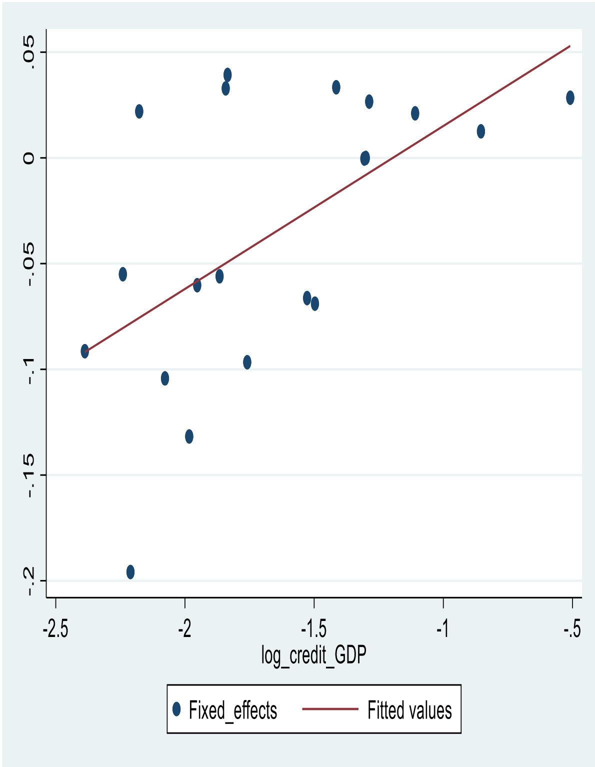
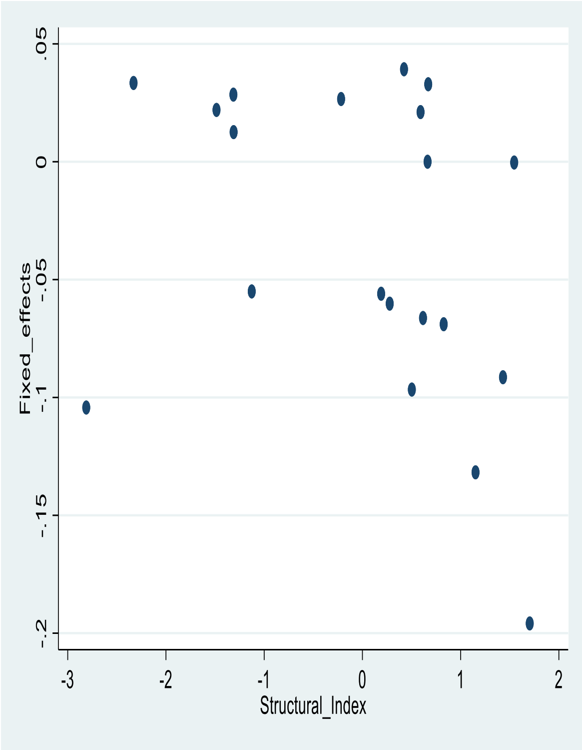


Figure 1C: Scatter plot between States Fixed Effect and log(Credit/GDP) ratio



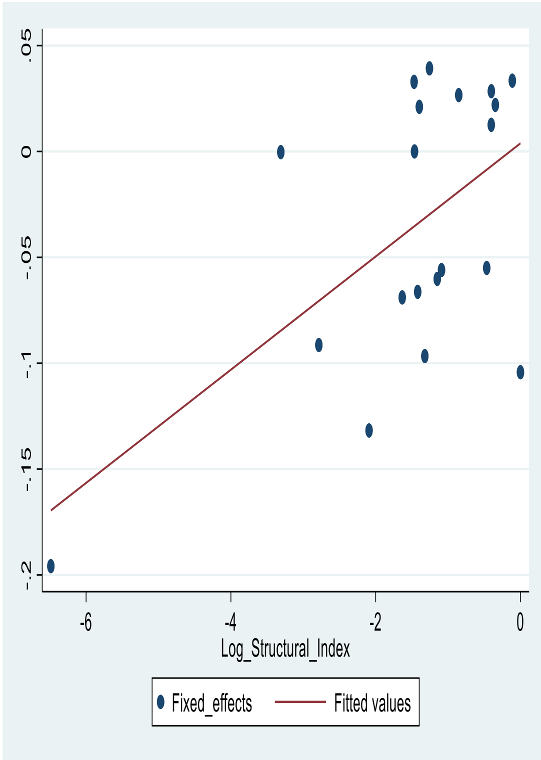
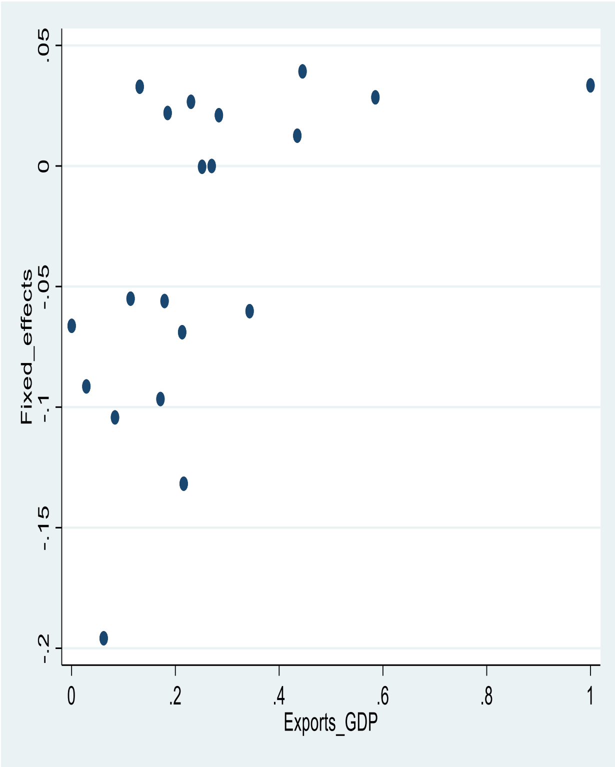


Figure 1D: Scatter plot between State Fixed Effects and Structural Index

Figure 1E: Scatter plot between State Fixed Effects and log of Structural Index.



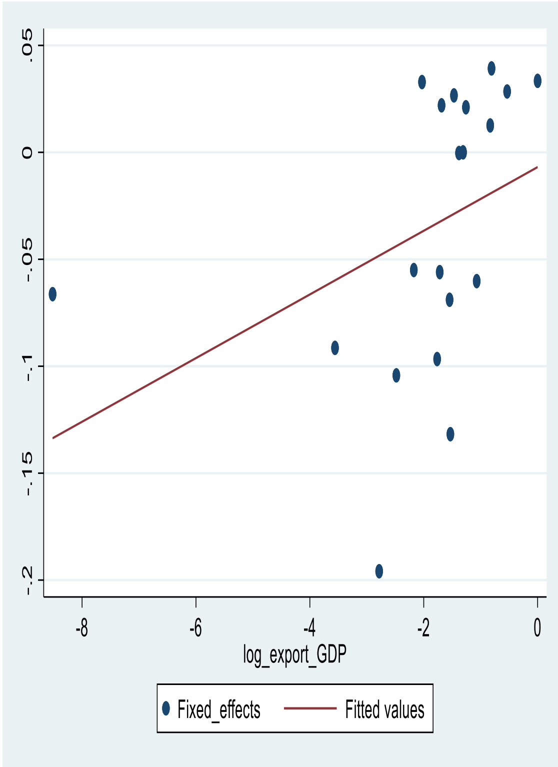


Figure 1F: Scatter plot between State Fixed Effects and Exports/GDP ratio

Figure 1G: Scatter plot between State Fixed Effects and log of Exports/GDP ratio

	Model A
Variables	Coefficient
Constant	0.03***
N	-0.55***
Log Credit/GDP Ratio	0.04**
Log Industry share/Agriculture share	0.05*
R-square- 0.60	
	Model B

Variables	Coefficient
Constant	0.13**
N	-0.51**
Log Credit/GDP Ratio	0.05**
Log Services share/Agriculture share	0.01
R-square- 0.45	

Table 1B: Regression Results of Model 1, Model 2, and Model 3

Note: *** denotes 1% significance level, **denotes 5% significance level and *denotes 10%significance level