Working Paper 268

Remoteness and Unbalanced Growth: Understanding Divergence Across Indian Districts

Samarjit Das
Chetan Ghate
Peter E. Robertson

September 2013

INDIAN COUNCIL FOR RESEARCH ON INTERNATIONAL ECONOMIC RELATIONS

Contents

Remoteness and Unbalanced Growth:
Understanding Divergence Across Indian Districts

Samarjit Das[ϒ], Chetan Ghate[±] and Peter E. Robertson[$]

Abstract

The existing literature on Indian growth finds no evidence of convergence across states. This represents a puzzle given the relatively free flows of capital, labor and commodities across state borders. We use a new data set of district level income and socio-economic data to explore the hypotheses of conditional convergence, using distance as an indicator of internal geographical trade and migration costs. We find evidence of conditional convergence for Indian districts but at a rate that is only half of Barro's "Iron Law". The results suggest that district level differences in trade and transport costs, infrastructure, and literacy rates, as well as state-level effects, have all contributed to the lack of absolute convergence in India.

JEL Classification*: O4, O5.*
Keywords*: Convergence, India, Economic Growth, Gravity Models.*

E-mail: *peter.robertson@uwa.edu.au*

[ϒ] Indian Statistical Institute, Kolkata
[±] Indian Statistical Institute, Delhi
[$] Corresponding Author: Peter Robertson, The Business School, University of Western Australia, M251 UWA, Crawley, Perth, W.A, Australia. Email: peter.robertson@uwa.edu.au. We thank Anu Rammohan, Ken Clements, Leandro Magnusson, M.H. Suryanarayana, Rod Tyers, participants at the Australian Development Economics Workshop, Monash University 2012, the 8th ISI Conference on Economic Growth and Development, ISI Delhi, 2012, and the University of Western Australia Economics Seminar Programme. We also would like to thank Shradha Mohta and Theo Backhouse for excellent research assistance. Robertson gratefully acknowledges funding from the Australian Research Council (DP110101316). Ghate and Robertson gratefully acknowledge funding from the Australia-India Institute. Ghate acknowledges hospitality from ICRIER during 2012-2013, where part of this project was completed.

1 Introduction

India's tentative economic miracle faces many hurdles, but one of the chief difficulties is sustaining the political impetus for reform. This is rendered more difficult by the fact that growth has been unbalanced – both across states and between urban and rural areas (Bardhan 2010). The growing regional disparities appear to have dampened political resolve for further economic reforms that might amplify inequality.

Understanding the cause of this unbalanced growth, and the factors that hasten or impede the benefits of economic reform, goes to the heart of the current debate over the role of markets and government in sustaining India's growth.[1] The existing literature on regional convergence in India, however, has largely been constrained to the analysis of inter-state differences. This literature finds little evidence of convergence. Rather, the pattern is one of divergence or, convergence to a bimodal distribution (Cashin and Sahay 1996, Rao and Sen 1997, Rao, Shand and Kalirajan 1999, Trivedi 2003, Bandopadhyay 2004, Ghate 2008, Kar, Jha and Kateja 2011, Das 2012, Ghate and Wright 2012, Bandopadhyay 2012).[2]

These findings for Indian States are curious since the hypothesis of absolute β–convergence has found widespread support in other countries (Sala-i Martin 1996, Durlauf, Johnson and Temple 2005, Barro 2012).[3] Moreover Desmet, Ghani, O'Connell and Rossi-Hansberg (2012) find that, along with its bias towards the service sector, India's spatial pattern of growth appears to show a much higher than usual difference in growth rates in large cities and medium sized urban areas.

In this paper we therefore aim to explain the pattern of Indian's growth at the district level, by using two new data sets on per capita incomes and social and economic characteristics for 575 districts. We find that the pattern of divergence remains when we analyze growth patterns at the district level. To try and explain this pattern we then draw on the economic geography literature that emphasizes trade and transport costs, and consider a model of conditional convergence across districts. Specifically we include the remoteness of each district, measured as the minimum distance to a major metropolitan center, as an explanatory variable, along with a range of district level characteristics.

[1]For a recent example see the debate between Bhagwati and Panagariya (2013) and Dreze and Sen (2013).

[2]Likewise there is evidence of growing inequality across India, such as Mishra and Kumar (2005), Chamarbagwala (2008) and Chaudhuri and Ravallion (2007).

[3]There is also no consensus on what the sources of this divergence might be. For example Cain, Hasan and Mitra (2012) find that states that are more open, with more roads and less labor market regulation, fared better. However Krishna and Sethupathy (2012) argue that the evidence of links between inequality and reforms in India are fairly weak.

Despite absolute divergence across districts, we do find strong evidence of conditional convergence across Indian districts, and that remoteness is a robust explanatory variable in explaining these differences. From a policy standpoint our analysis points to the importance of infrastructure differences, as indicated by urbanization and electricity, and state level policy differences, as important variables in understanding differences in regional incomes and growth rates. Nevertheless while the results do shed light on the pattern of absolute divergence, the rate of β–convergence of approximately -1.3%, is only half the value of Barro's "Iron Law" of convergence, (Barro 2012). This suggests that other significant regional economic barriers to convergence exist in India.

2 Preliminary Statistical Analysis

2.1 The Data set

To investigate the pattern of growth across India we use two new data sets of district level incomes and social and economic characteristics – respectively the *Indicus* "Development Landscape" and "District GDP" data-sets. The data consist of 575 district level observations of district income for two years, 2001 and 2008.[4]

The availability of district level income data provides the opportunity to observe regional disparities in India at a much finer level than previous studies based in Indian States. This is also advantageous insofar as there is likely to be a larger degree of heterogeneity in income levels, growth rates and other characteristics such as urbanization or literacy, compared to state level data.

We begin with a preliminary exploration of the data by considering different indicators of convergence and how the shape of the distribution of district incomes has changed over time. First Table 1 shows the wide disparity in income levels across states. There is a 9.8 fold difference between the richest state *Goa*, and the poorest state *Bihar*. This is larger than the real income gap between the GDP per capita of the USA and Angola, and only slightly smaller than the real income gap between the USA and India.[5]

At the district level, however, that gap is much larger. The range in per capita incomes in 2008 is from a minimum of RS. (m) 3858 in the *Sheohar* district (*Bihar*) to a maximum of

[4] This data has attracted some debate. See Himanshu (2009) but also, importantly, the reply by Bhandari (2009).

[5] This comparison is based on the Penn World Tables PPP values, that report Angola with a relative per capita GDP of 11.51 and India 7.21 in 2008.

RS. (m) 139868 in *Jamnagar* (*Gujarat*) This is an income ratio of 36 which is equivalent, for example, to the ratio between the USA and Rwanda according to the Penn World Tables.

The district data are shown visually in Figure 1. It can be seen that there are generally lower incomes in central districts, particularly in the eastern states. Likewise the wealthy western corridor running from north of Delhi down the west coast through Western *Maharashtra* and *Karnataka*, *Goa* and *Kerala* is easily observed. Figure 1 is thus suggestive of a strong geographic pattern in the differences in incomes across India.

The fact that the within-India differences are comparable to cross-country differences is remarkable given that there are no political barriers to migration, approximately free trade, and a common set of federal institutions, policies and governance. That such differences could persist over time is in stark contradiction to the standard competitive model that motivates the extensive literature on absolute β–convergence across regions. In contrast, it points to the potential relevance of trade barriers, transport costs and conglomeration effects as emphasized in the economic geography literature.

2.2 Absolute β and σ convergence

The standard concept of convergence in cross sectional regional data is absolute β–convergence (Baumol 1986, Sala-i Martin 1997, Durlauf et al. 2005). This is given by the coefficient β from (1)

$$y_{i,t} - y_{i,0} = \beta \, y_{i,0} \; + \; \varepsilon_i \tag{1}$$

where $y_{i,t}$ is the natural log of income at time t in region i and $y_{i,0}$ is initial income.[6] The left hand side of (1) represents the growth rate over the period $(0, t)$. The results of estimating (1) across Indian districts are given in Table 1. It can be seen that across India there is strong evidence of a small rate of divergence with β =0.007, which is statistically significant at the 1 % level. Hence, on average, richer districts have been growing slightly faster than poorer districts.

Table 1 also shows the results of estimating (1) for each state separately. Thus we ask whether there is convergence across districts within each state. In four states, *Assam*, *Chhattisgarh*, *Kerala* and *Rajasthan*, there is significant absolute β–convergence of district level incomes. However there is also significant within-state divergence in three states –

[6]We report β for all states except *Goa*, *Pondichery* and *Chandigarh* where the number of districts is 2 or 1.

Haryana, *Orissa* and *Uttar Pradesh (UP)*.[7] Nevertheless for the vast majority of states the estimated β – convergence coefficient is insignificantly different from zero. Thus there is little evidence of strong convergence, either across the country as a whole or within individual states.

Next we consider σ – convergence, which is defined as a decline in the variance of district level per capita log incomes across time. Table 2 shows the variance of district log incomes in the two periods, 2001 and 2008. It can be seen that there was a 30.7% increase in the variance of log incomes across districts – from 0.27 to 0.35. Thus there has also been σ – divergence.

Table 2 reports a simple variance decomposition using log per capita incomes.[8] Here, *within-state* variance, ν^W refers to deviations of district log incomes, y_{ij}, from their state level mean log income, $\bar{y}_j$, $y_{ij} - \bar{y}_j$, and *between-state* variance, ν^B, refers to deviations of state level mean log incomes $\bar{y}_i$ from the country-wide mean log income, $\bar{y}$, $\bar{y}_i - \bar{y}$. By definition the total India-wide variance of incomes across all districts, ν^T, is equal to the sum of the within-state variance and between state variance, $\nu^T = \nu^W + \nu^B$. This variance decomposition shows that there has been a similar increase in σ–divergence both within states and between states.

Further evidence on the pattern of Indian growth can be obtained by examining other aspects in the change in the distribution of district incomes. To that end Figure 2 plots the kernel density estimate of the probability density function (PDF) for district log incomes for 2001 and 2008.

It shows the shift in mean income; a fall in peakedness (kurtosis) with a slight increase in concentration on the left tail (skewness). Likewise the Cumulative Distribution Function (CDF) in Figure 3 shows that each district has become better off in 2008 as compared to 2001. Together these visual images suggest while the income distribution has widened at the upper tail, incomes have increased at each point on the distribution. There is significant churning within the distribution, and only sixteen districts remain in the same position on the distribution between 2001 and 2008. Overall however Kendall's rank correlation *tau* statistic is 0.8, suggesting a high correlation of rankings between the two periods.

Thus, though there is some evidence of convergence within a few states, among most states there is no correlation between initial income and growth. Examining the country

[7] Moreover both *UP* and *Orissa* are among the poorest states with the largest primary sector income shares, above 30%.

[8] Details of this simple decomposition are given in the appendix.

as a whole, there is evidence of β and σ divergence, reflecting faster growth in higher income districts with most districts experiencing growth across the entire distribution.

3 Conditional Convergence and Geography

The preceding model of absolute β–convergence explicitly assumes that all regions within a country have the same steady state income level (Barro and Sala-i Martin 1991, Durlauf et al. 2005, Barro and Sala-i Martin 2005). This can be justified, for example, by the factor price equalization theorem, which states that free-trade and identical technologies will result in a convergence of incomes across regions. More generally factor mobility will result in absolute convergence, even in the absence of identical technologies.

However the economic geography literature following Krugman (1991) and Lucas (1988) has emphasized the importance of barriers to trade, information or factor migration, and regional externalizes and conglomeration effects. Thus, even in a regional context, there may be significant obstacles to convergence and hence long run differences in per capita incomes.[9]

In view of this, the concept of conditional convergence may be appropriate. Specifically, consider a long run equilibrium where all districts are growing at rate g. Denote productivity at time t, measured in effective labor units, as $A_i(t)$ and assume that $A_i(t) = A_i(0)^{gt}$. Then on a balanced growth path, district income per effective worker $\hat{y}_i^* \equiv (y_i^*/A_i^*)$ will be a constant.

Next suppose that the convergence path to the steady state, or balanced path equilibrium is given by a standard partial adjustment model

$$\ln \hat{y}_i(t) - \ln \hat{y}_i(0) = \beta(\ln \hat{y}_i^* - \ln \hat{y}_i(0)), \tag{2}$$

which says that the current growth rate of district i depends on the gap between the current income level and the long run balanced path level, both measured in terms of output per effective worker.

In a regional setting, however, the concept of conditional convergence presupposes some significant regional economic barriers, such as barriers to trade and migration. Thus, though the form (2) is familiar from the cross country literature, in our regional context we also need a theory of how differences in regional steady state incomes, y_i, arise.

[9] In a less formal way these characteristics also featured in earlier development literature such as Lewis (1955).

As noted above the economic geography literature emphasizes the geographic costs of migration and transport and the role of cities including the fact that in developing economies most migration is from rural districts to urban centers. In particular Ciccone and Hall (1996) and Glaeser, Kallal, Scheinkman and Shleifer (1992) note that much of the growth process, such as technology adoption and capital accumulation, occurs in cities, and that high density clusters are important sources of growth. With respect to India specifically, Desmet et al. (2012) have argued that growth is particularly concentrated in the largest cities.

As in the trade-gravity literature we can capture the degree of trade, transport and migration costs by using measures of distance (Anderson and Wincoop 2004). Specifically, a district that is very remote from any large metropolitan center might have a low long-run per capita income level, relative to one that is very close to the same major city.

We can formalize this idea as follows. Let y_i denote district i per capita income and y^* denote the steady state income per worker in a nearby metropolitan center. Then for district i we may consider a variable θ_i such that, in a steady-state equilibrium,

$$y_i^* = \theta_i \, y^* \tag{3}$$

where y_i^* is the steady state income per person for district i and θ_i measures the extent of all barriers to complete convergence, such as trade and transport costs, communications costs, road quality and other geographic barriers. The variable θ_i thus determines the maximum degree of convergence, or catch-up, that can be obtained. Specifically if $\theta_i < 1$ district i will only achieve partial convergence to the metropolitan center.

In terms of effective workers (3) implies $\hat{y}_i^* = \theta_i \; \hat{y}^*$.[10] Then using (2) the transitional growth process for some non-metropolitan district i, can be derived as

$$\ln y_i(t) - \ln y_i(0) = gt - \beta \ln y_i(0) + \ln A_i(0) + \beta \left(\ln \hat{y}^* + \ln \theta_i \; \right). \tag{4}$$

In equation (4) the growth rate of district i depends on: (i) the initial per capita income of district i, $y_i(0)$; (ii) the level of labor productivity of district i, $A_i(0)$; (iii) the steady-state value of income per effective worker in the relevant metropolitan center, $\hat{y}^*$ and; (iv) the distance between district i and the metropolitan center, θ_i.

Thus (4) says that the growth rate of per capita income for some district i, depends on

[10] We assume long run technology convergence so that $A_i^* = A^*$. Alternatively one could assume that technological gaps exist in the long run and that this difference is absorbed as an argument in the function θ_i.

its current level of per capita income relative to its long run balanced growth path level which, in turn, depends in part on the remoteness of the district from the metropolitan center and its steady-state per capita income level. We can use this to explain the growth and convergence of the non-metropolitan districts, taking the growth of the metropolitan centers as given. The aim therefore is not to explain India's overall growth rate, since we explicitly take the growth of key centers as different. Rather we consider the extent to which various non-metropolitan districts are sharing in the growth process.

To implement equation (4) empirically we first need to define what we mean by a "metropolitan center". As shown in Table 3, India has three mega-cities with populations above 10 million, *Delhi*, *Mumbai*, and *Kolkata*. Of these, *Delhi* and *Mumbai* have extended urban agglomerations – defined as areas of unbroken urbanization – that exceed 20 million. Nevertheless even the smaller cities, *Bangalore*, *Hyderabad* and *Ahmedabad*, have populations of over 6 million and there are ten Indian cities with urban agglomerations over 3 million. We begin therefore by initially defining the "metropolitan centers" as the seven largest Indian cities, which include all cities that had populations over 6 million. As a robustness check we also consider alternative definitions up to the ten largest cities listed in Table 3. As we shall see, the results are very robust to these alternative definitions.

Next we define the variable *Distance*, D_i, as the minimum distance, by road, between district i and the closest metropolitan center.[11] Figure 4 shows D_i for each district in India. Given the location of the seven largest cities the map shows a band of relatively remote districts between *Delhi* and *Hyderabad* through *Madhya Pradesh* and *Chhattisgarh*. The remaining remote districts are located in the geographic extremities, especially the far north of *Jammu and Kashmir*, the eastern most districts of *Gujarat* and the far western districts. It can also be seen that there are clusters of less remote districts along the western corridor from *Delhi* to *Bangalore* and *Chennai*. [12]

The final step needed to operationalize (4) is to specify an empirical counterpart to (3), which is a function of the minimum distance from a district to a metropolitan center. The gravity literature in international trade suggests a simple inverse relationship such as $\theta_{i,j} = \theta\ D_i^{\gamma}$. Hence, using logarithms we have

$$\ln \theta_{i,j} = \ln \theta\ + \gamma\ \ln D_i + \eta\, \mathbf{X}_i \qquad (5)$$

[11]The data on distance between districts is from *Google Maps* and a variety of other sources including Indian state tourism data. It denotes the minimum distance (by road) from one district headquarters to another.

[12]This picture of a western corridor of relative urbanization is even stronger if we move to consider the ten largest metropolitan centers.

where $\gamma < 0$, is the distance elasticity, $\mathbf{X}_i$ is a vector of characteristics of region i and η is a vector of coefficients. This follows the standard gravity model, familiar in the trade literature.[13]

From (4) and (5) we obtain an empirical model,

$$\ln y_i(t) - \ln y_i(0) = \alpha_0 + \alpha_1 \ \ln y_i(0) \ + \ \alpha_2 \ \ln D_i + \eta \, \mathbf{X}_i \ + \ \epsilon_i \qquad (6)$$

where $\alpha_1 \equiv -\beta$, $\alpha_2 = \beta\gamma$, $\alpha_0 = g + \beta \ln A_i(0) + \beta \ln \hat{y}^* + \theta$, and $\ln A_i(0) = \ln A + \epsilon_i$, where ϵ_i is a district specific random shock reflecting, for example, institutions, climate and endowments.

Equation (6) is our base-line model. The convergence coefficient captures the notion that the larger the gap between the i^{th} district and the metropolitan center in the initial time period, the lower the growth rate. Distance is expected to negatively affect district incomes relative to the closest metropolitan center and hence reduces the transitional growth rate.

Finally, a further simple extension of (6) allows for the possibility that the metropolitan districts have different balanced path income levels. Specifically, suppose $\hat{y}_j^* = f(\mathbf{Z}_j)\ \hat{y}^*$, where $\mathbf{Z}_j$ is a vector of characteristics that affect the steady state income levels of metropolitan center j. Then, assuming $f(\mathbf{Z})$ is log linear gives

$$\ln y_i(t) - \ln y_i(0) = \alpha_0 - \alpha_1 \ln y_i(0) \ + \ \alpha_2 \ \ln D_i + \eta \, \mathbf{X}_i + \delta \, \mathbf{Z}_j + \epsilon_i. \qquad (7)$$

In what follows we estimate (6) and (7) using our cross-section of Indian districts.

4 Results

As discussed above, our data consists of district level GDP growth rates and district level characteristics from the *Indicus* data sets. Summary statistics for the key variables of interest are given in Table 4.[14]

[13]This also requires the restriction that $D_i \geq 1$, which will be true in our data.

[14]A visual inspection of the data suggests the presence of heteroscedasticity and the Breusch-Pagan (BP) test for heteroscedasticity on preliminary OLS results confirms this. As the form of heteroscedasticity is unknown, the application of GLS is not feasible. The implication of heteroscedasticity is that OLS will result in biased standard errors and tests based on these standard errors will be invalid. In what follows we therefore use White's (1982) robust standard errors to obtain valid inferences, even though efficiency is sacrificed.

The results for our baseline model, equation (6), are given in Table 5a. The conditioning vector consists of the percentage of households with *electricity*, the number of *commercial banks*, *urbanization,* and the percentage of *irrigated land*. We also include state dummy variables and report the results of an F-test for the joint significance of these state dummy variables. Columns (1)-(7) include our remoteness variable *Distance* which, as discussed above, measures the minimum road distance between each district head quarters to each of India's seven biggest metros (*Delhi*, *Mumbai*, *Kolkata*, *Chennai*, *Bangalore*, *Hyderabad*, and *Ahmedabad*). Column (8) includes all of the conditioning variables except *Distance*.

It can be seen that the sign of the convergence coefficient β, is significant and negative across all models. Thus allowing for different district characteristics has overturned the finding of divergence across districts to one of conditional convergence. This is important since it suggests that the observed pattern of divergence can be understood as resulting from different long run steady state income levels of each district which are explained by the conditioning variables in the model. In particular, it can be seen that *Distance* is significant to at least the 5% level in all regressions. As expected, an increase in *Distance* reduces steady-state income level and hence also reduces the transitional growth rate for a given level of initial income, $y(0)$. Furthermore, across all models it can be seen that the variables *urbanization* and *electricity* are significant at the 1% level. Likewise *irrigated land* is significant.[15]

Arguably since electricity is government controlled the significance of these suggests that differences in public infrastructure and local governance are important in understanding differences across districts. Moreover the F-tests for the joint significance of the state dummy variables is also very significant across the various models. We discuss these results further below.

Though we have found strong evidence of conditional convergence, the estimated value of β = -0.8% to 1.29% is much slower than the values found in the growth literature using quite different regional aggregations across a wide array of counties. In particular, it is roughly half of Barro's "iron law of convergence" (Sala-i Martin 1996, Sala-i Martin 1997, Barro 2012). Even the largest rate estimated, of approximately -1.3% (Column 7, Table 5a) implies that the gap between each district's current income level, and its long run or steady state income level, is halved only every 62 years. Thus, though we find convergence, it implies very little actual convergence would occur over typical policy horizons. For example, at this rate, at the end of a decade a per capita income gap between two districts would still be 90 percent of the gap that existed at the start of the

[15]The sign in this case is negative suggesting that high land productivity reduces migration.

decade.

In Table 5b, we allow for the possibility of different metros having different steady-state incomes, as in (7). Thus, in addition to the RHS explanatory variables in Table 5a, we include the $\mathbf{Z}_j$ vector of characteristics that affect the steady state income levels of metropolitan center j. It can be seen, however, that across the various specifications in Table 5b, the additional explanatory variables are generally insignificant with the exception of *The Metropolitan Districts Literacy rate*, which tends to be significant to at least the 10% level across most specifications. This is interesting since district literacy was not found to be significant. Thus the results suggest that literacy is an important fact in explaining differences in growth rates across Metropolitan areas, but not across the country more generally. Again this concords with theories of growth and urban agglomerations which emphasizes complementarities between network externalizes and human capital.

Including the additional metropolitan characteristics $\mathbf{Z}_j$ however has little effect on the estimated convergence and distance characteristics, which tend to be very stable across all specifications in Table 5a and 5b. Hence the conclusions on the low rate of convergence are robust to the inclusion of the metropolitan characteristics. Likewise *Distance* continues to be statistically significant in all specifications of the models.

4.1 The Impact of Remoteness

To what extent do different degrees of remoteness matter for understanding differences in growth and incomes across India. The elasticity of *Distance* with respect to steady state income is given by $\gamma = -\alpha_2/\alpha_1$. This value is reported for each model in Tables 5a and 5b, along with a joint significance test. It can be seen that the estimates of γ are significant at the 1% level across each model with a value ranging from approximately -0.57 to -0.25.

To interpret this, consider two districts i and j with identical characteristics except for their distance from the metropolitan center k. Then from (6) we have

$$\frac{y_i^*}{y_j^*} = \left(\frac{D_{i,k}}{D_{j,k}}\right)^{\gamma}. \tag{8}$$

If, for example, the more isolated district, i, is twice the distance from the metropolitan center than the closer district, j, then from (8), we have $D_{i,k}/D_{j,k} = 2$.[16] Assuming the

[16]The mean distance is 532km with a standard deviation of just under 400km, so doubling the distance

most conservative estimate of the gravity parameter of $\gamma = -.25$ from Column 7, Table 5b, this implies $y_i^*/y_j^* = 2^\gamma = 0.84$. Thus our estimates imply that the more remote district will have a steady state income level that is approximately 84% of the closer district.[17]

The most remote district in our data is *Tamenglong*, in Manipur, which is a mountainous district near the Burmese border and is 2531 kilometers from *Kolkata*, the nearest metropolitan center. At the other end of the spectrum the district *South 24 Parganas* is only 7.9 kilometers from *Kolkata*. This gives a ratio of approximately 320 which means, from (8), that other thing equal we would expect the more remote district to have an income level of only 24% of the closer district. Thus the distance coefficient suggests quite a large impact on income levels for very remote districts but relative modest effects for districts that are within a range of twice or half the average distance.

In terms of growth rates the coefficient on distance $\alpha_2 = \beta\gamma$ ranges from approximately -0.003 to -0.005. This value is the partial effect of a 1 percentage point change in distance on the growth rate. Hence the estimates imply that a district that is twice as remote will have a transitional growth rate that is 0.20 to 0.35 percentage points lower than the closer district.[18] However at the maximum distance in the data, of 320 times, the more remote district would have a growth rate that is 1.7 to 2.9 percentage points lower. Thus the distance variable has an economically important effect on observed transitional growth rates for the very remote regions.

4.2 Discussion

The significance of the variables in Tables 5a and 5b sheds some light on the observed pattern of divergence across India. First Table 5a shows that the divergence of growth rates can be understood as resulting from differences in long run income levels. Our analysis points to the public infrastructure variable, as indicated by electrification, and urbanization being important determinants of long run district income levels, which is consistent with the recent study by Desmet et al. (2012) who note very strong agglomeration effects in India.

This is a useful starting point in considering potential policy responses to address the unbalanced nature of India's growth. It suggests that absolute convergence will depend

is just a little more than increasing the distance by one standard deviation from the mean.

[17]Likewise if the more remote district were to have have a steady-state income that is approximately 50% of the closer districts, it would need to be 16 times further from the center.

[18]Since $\alpha_2 = \beta\gamma = \partial g/\partial lnD$ and $\ln 2 = 0.69$.

on increasing equality in these conditioning variables. Likewise divergence in growth rates may be mitigated through improving economic policy, particularly infrastructure investment, in low growth regions. The potential policy role is underscored by the fact that the state dummy variables – which proxy for differences in state steady states – are highly significant. This is consistent with the literature cited in the introduction that has pointed to significant policy differences at the state level, particularly with respect to labor laws, (Besley and Burgess 2004, Acharya, Baghai and Subramanian 2010) Finally, we have also seen that remoteness is significant across all our models, which supports our conjecture that transport and information costs impose important regional constraints on development. The results suggest that this effect is particularly important in understanding the reasons for low growth rates in very remote districts, but only has a modest impact on most districts.

5 Robustness

5.1 Stability

As a robustness test we then extend our definition of a metropolitan center to include the 10 largest urban agglomerations in India by population as in Table 3.[19] The overall conclusion is also robust to these alternative definitions of distance or remoteness with very little change in significance of the key variables or the estimated size of the coefficients.[20] Second we consider whether our distance variable is stable across different data sets. To do this we divided the whole data set into several subgroups, and then examine stability of model parameters. To this end, we re-estimate (6) and (7) but drop several districts. Specifically, we first drop all north-east districts, then all districts from *Bihar* and *Maharashtra*. Other alternatives are given in Table 6.

A stability test is then conducted by using interaction dummy variables, where the dummy variable takes the value 1 for included districts and takes value 0 for excluded districts. Then we examine whether such interaction dummies are significant or not based on an F-test. The results are depicted in Table 6.

All the parameters, including the distance variable, were found to be very stable across the data subsets, as shown in Table 6 where the estimated p-values for the F-tests are

[19] Using data for the 10 biggest metros in India, the gravity parameter (controlling for metro steady state characteristics) is -.2417, very close to -.2509 reported in Table 5b. Because of space constraints, we do not include these results. These results are available from the authors on request.

[20] These results are available upon request.

significantly larger than 0.05. Thus we do not reject the null hypothesis of constant coefficients. Hence this test indicates there is no evidence that the parameters change across the subsets of the data districts.[21]

5.2 Endogeneity

Aside from these robustness tests we also consider the potential for the explanatory variables to be endogenous, leading OLS estimates to be biased and inconsistent. To investigate this we first apply the Hausman test by comparing 2SLS and the OLS estimates.[22] The Hausman tests are negative for all these cases, which is not unexpected since, as discussed above, there is evidence that our data are strongly heteroscedastic, invalidating the use of the Hausman test.

We therefore compare the equality of two parameter vectors (OLS and 2SLS) in a SUR setting. Table 7 provides results of this endogeneity test for three variables: *District Literacy*; *District Pucca Roads*, and; *District Urbanization*, as included in Tables 5a and 5b.[23] The test statistic follows a χ^2 distribution with the number of model parameters as the degrees of freedom. [24] We find that we cannot reject the null hypothesis of no endogeneity.

6 Conclusion

India's growth has been very unbalanced, with growing inequality, a bias towards services and an excessive concentration of its growth in a few large cities and divergence of incomes across states. The causes of this pattern of divergence - in this regional context with free trade and factor mobility - are, however, not well understood.

We therefore examine the evidence for convergence of per-capita incomes at the district level using a new data set of district incomes and socio-economic characteristics. We find little evidence of convergence either within states or across all districts as a whole.

[21] We examine parameter stability for the genuine regressors excluding the intercept and the state dummy variables. Note also that it is important that these subsets of the full data set are selected in a random fashion. For example creating subsets of the data based on different income groups would introduce a sample selection problem.

[22] For 2SLS the identifying variables we use are the percentage of household with telephones, percentage of people below the poverty line and female literacy rates

[23] We have also tested exogeneity status of the *Metro* variables included in Tables 5a and 5b. The SUR framework based tests strongly accept the null hypothesis of exogeneity.

[24] For example, Model (iii) has 38 parameters and Model (iv) has 39 parameters.

Rather there is β–divergence across all districts and also an increase in the variance of log incomes across districts over time, or σ – divergence.

To better understand these facts we then consider a model of conditional β–convergence across Indian districts. In order to capture the implied trade and transport costs in a model of regional convergence, we include the distance between each district and the closest large metropolitan centers as a conditioning variable. We find strong evidence of conditional convergence between Indian districts. The key explanatory variables are urbanization and electrification in the non-major metropolitan districts. Thus the results support Desmet et al. (2012) who argue that frictions, policies, and a general lack of infrastructure in medium-density cities is preventing the spread of growth in India. Likewise the differences across states support studies that have emphasized the role of different degrees of regulation across states (Besley and Burgess 2004, Acharya et al. 2010). We also find some evidence that literacy rates are important in explaining differences in income levels across the metropolitan centers.

Though we find evidence of conditional convergence, the rate of approximately -1.1 to 1.3%, is only half of Barro's "iron-law". Thus conditional convergence, though significant, is nevertheless a very weak force. Likewise remoteness only has a very large impact on the most remote districts. Hence the reasons for the lack of rapid convergence of district incomes remain uncertain. Nevertheless we have made some progress in impotent conditioning factors and also in identifying a potential role for public investment programmes in addressing some of the imbalances in India's growth.

References

Acharya, Viral V., Ramin P. Baghai, and Krishnamurthy V. Subramanian (2010) 'Labor laws and innovation.' Technical Report 16484, NBER Working Paper

Anderson, James E., and Eric van Wincoop (2004) 'Trade costs.' *Journal of Economic Literature* 42(3), 691–751

Bandopadhyay, S. (2004) 'Twin peaks-distribution dynamics of economic growth across indian states.' In *Growth, inequality and poverty: prospects for pro-poor growth,* ed. A. Shorrocks and R. van der Hoeven (New York: Oxford University Press)

—— (2012) 'Convergence club empirics: Evidence from indian states.' *Research in Economic Inequality (forthcoming)*

Bardhan, Pranab (2010) *Awakening Giants: Feet of Clay* (Princeton N.J.: Princeton University Press)

Barro, R.J., and X Sala-i Martin (1991) 'Convergence Across States and Regions.' *Brookings Papers on Economic Activity* 1(1), 107–182

Barro, Robert J. (2012) 'Convergence and modernization revisted.' *Paper Presented at the Nobel Symposium on Growth and Development, Stockholm, September 3-5, 2012*

Barro, Robert J., and J. Sala-i Martin (2005) *Economic Growth* (Singapore: McGraw Hill)

Baumol, William J. (1986) 'Productivity growth, convergence, and welfare: What the long run data show.' *American Economic Review* 76(5), 1027–1085

Besley, Timothy, and Robin Burgess (2004) 'Can labor regulation hinder economic performance? evidence from india.' *The Quarterly Journal of Economics* 100(6), 91–134

Bhagwati, Jagdish, and Arvind Panagariya (2013) *Why Growth Matters: How Economic Growth in India Reduced Poverty and the Lessons for Other Developing Countries* (PublicAffairs)

Bhandari, Laveesh (2009) 'Socio-economic performance of constituencies: A response.' *Economic and Political Weekly* 44(40), 61–63

Cain, J, Rana Hasan, and Devashish Mitra (2012) 'Trade liberalization and poverty reduction: New evidence from indian states.' In *India's Reforms: How They Produced*

Inclusive Growth, ed. Jagdish Bhagwati and Arvind Panagariya (N.Y.: Oxford University Press) chapter 4, pp. 91–169

Cashin, Paul, and Ratna Sahay (1996) 'Internal migration, center-state grants, and economic growth in the states of india.' *International Monetary Fund Staff Papers* pp. 123–171

Chamarbagwala, Rubiana (2008) 'Economic liberalization and wage inequality in india.' *World Development* 34(12), 1997–2015

Chaudhuri, Shubham, and Martin Ravallion (2007) 'Partially awakened giants: Uneven growth in china and india.' In *Dancing With Giants: China, India, And The Global Economy,* ed. L. A. Winters and Shahid Yusuf (Washington State DC: The World Bank)

Ciccone, Antonio, and Robert E Hall (1996) 'Productivity and the density of economic activity.' Technical Report, National Bureau of Economic Research

Das, Samarjit (2012) 'The convergence debate and econometric approaches: Evidence from india.' In *The Oxford Handbook of the Indian Economy,* ed. C. Ghate (New York: Oxford University Press) pp. 766–782

Desmet, Klaus, Ejaz Ghani, Stephen O'Connell, and Esteban Rossi-Hansberg (2012) 'The spatial development of india.' Technical Report 6060, World Bank Policy Research working paper

Dreze, Jean, and Amartya Sen (2013) *An Uncertain Glory: India and its Contradictions* (Allen Lane)

Durlauf, S. N., P. A. Johnson, and J. R. W. Temple (2005) 'Growth econometrics.' In *Handbook of Economic Growth,* ed. P. Aghion and Durlauf S. N., vol. 1A (North-Holland: Amsterdam) pp. 555–677

Ghate, Chetan (2008) 'Understanding divergence in india: A political economy approach.' *Journal of Economic Policy Reform* 11(1), 1–9

Ghate, Chetan, and Stephen Wright (2012) 'The v-factor: Distribution, timing and correlates of the great indian growth turnaround.' *Journal of Development Economics* 99(1), 58–67

Glaeser, Edward L, Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer (1992) 'Growth in cities.' *Journal of Political Economy* 100(6), 1126–52

Government of India (2013) 'Press information bureau.' http://pibmumbai.gov.in/scripts/detail.asp?releaseId=E2011IS3

Himanshu (2009) 'Electoral politics and the manipulation of statistics.' *Economic and Political Weekly* 44(19), 31–35

Kar, Sabyasachi, Debajit Jha, and Alpana Kateja (2011) 'Club-convergence and polarization of states: A nonparametric analysis of post-reform india.' *Indian Growth and Development Review* 4(1), 53 – 72

Krishna, Pravin, and Guru Sethupathy (2012) 'Trade and inequality in india.' In *India's Reforms: How They Produced Inclusive Growth,* ed. Jagdish Bhagwati and Arvind Panagariya (N.Y.: Oxford University Press) chapter 6, pp. 247–278

Krugman, Paul R. (1991) 'Increasing returns and economic geography.' *Journal of Political Economy* 99(1), 483–499

Lewis, Arthur (1955) *The Theory of Economic Growth* (Allen and Unwin)

Lucas, Robert E. (1988) 'On the mechanics of economic growth.' *Journal of Monetary Economics* 22(1), 3–42

Mishra, Prachi, and Utsav Kumar (2005) 'Trade liberalization and wage inequality: Evidence from india.' *IMF Working Paper WP0520*

Rao, G., R. Shand, and K. Kalirajan (1999) 'Convergence of incomes across indian states: A divergent view.' *Economic and Political Weekly* 17(13), 769–778

Rao, M Govinda, and Kunal Sen (1997) 'Internal migration, center-state grants, and economic growth in the states of india: A comment on Cashin and Sahay.' *Staff Papers-International Monetary Fund* 44(2), 283–288

Sala-i Martin, Xavier (1997) 'I just ran two million regressions.' *American Economic Review* 87(2), 178–83

Sala-i Martin, Xavier X. (1996) 'Regional cohesion: Evidence and theories of regional growth and convergence.' *European Economic Review* 40(6), 1325–1352

Trivedi, K. (2003) 'Regional convergence and catch-up in india between 1960 and 1992.' *Mimeo, Oxford University Economics Department*

Table 1: Within State Convergence

State	Pop (Millions)	Per Capita GDP Rs 000's 2007-08	Share Primary Sector %	β	p-value
All India	1,137.1	38	21	0.0107***	(0.0019)
Andhra Pradesh	82.2	38	29	-0.0032	(0.0069)
Arunachal Pradesh	1.2	34	26	-0.0134	(0.0345)
Assam	29.3	24	35	-0.0332***	(0.0091)
Bihar	95.6	11	25	-0.0068	(0.0138)
Chhattisgarh	23.2	33	24	0.0188**	(0.0080)
Gujarat	55.9	52	19	0.0012	(0.0057)
Haryana	23.8	62	21	0.0333*	(0.0114)
Himachal Pradesh	6.5	49	22	0.0081	(0.0308)
Jammu and Kashmir	11.0	29	27	0.0047	(0.0098)
Jharkhand	30.2	23	22	0.0304	(0.0179)
Karnataka	56.7	38	19	0.0102	(0.0091)
Kerala	33.8	48	17	-0.0391*	(0.0206)
Madhya Pradesh	69.0	20	33	-0.0005	(0.0096)
Maharashtra	107.1	53	13	0.0119*	(0.0065)
Manipur	2.4	24	26	-0.0009	(0.0184)
Meghalaya	2.5	30	27	0.0102	(0.0164)
Mizoram	1.0	34	15	0.0176	(0.0130)
Nagaland	2.2	33	34	-0.0157	(0.0305)
Orissa	39.7	26	31	0.0492***	(0.0085)
Punjab	26.4	52	31	-0.0054	(0.0298)
Rajasthan	64.1	26	28	-0.0338***	(0.0123)
Tamil Nadu	66.0	44	14	0.0089	(0.0092)
Uttar Pradesh	189.3	18	31	0.0133***	(0.0046)
Uttaranchal	9.4	36	20	0.0080	(0.0147)
West Bengal	86.4	35	23	0.0033	(0.065)

Note 1: *, **, *** denotes 10, 5 and 1 percent levels of significance respectively.

Note 2: Robust (White) standard errors are used.

Table 2: Decomposition of σ–Convergence

	Variance	Between State Variance	Within State Variance	Skewness	Kurtosis	Gini
2001	0.27	0.15	0.12	0.15	3.09	0.0307
2008	0.35	0.20	0.15	0.16	2.88	0.0322
Change	0.08	0.05	0.03			

Table 3: Metropolitan Districts

Extended Urban Agglomeration	Population 2011 (Millions)
Delhi	21,753,486
Greater Mumbai	20,748,395
Kolkata	14,617,882
Chennai	8,917,749
Bangalore	8,728,906
Hyderabad	7,749,334
Ahmedabad	6,352,254
Pune	5,049,968
Surat	4,585,367
Jaipur	3,073,350

Source: Government of India (2013)

Table 4: Descriptive Statistics

	Mean	Variance	Minimum	Maximum	Skewness
Per capita GDP	9.583	0.274	8.243	11.313	0.148
Distance	6.004	0.671	2.067	8.018	-1.091
Literacy	4.131	0.046	3.408	4.570	-0.750
Electricity (%)	3.776	0.578	1.131	4.588	-1.212
Commercial Banks	-9.698	0.175	-11.194	-8.227	0.500
Urbanization	2.870	0.565	0.279	4.605	-0.199
Irrigated Land	-3.253	1.163	-7.782	-1.139	-0.980
Pucca Road	3.968	0.617	-1.204	4.605	-3.063
Metro Electricity	4.557	0.000	4.543	4.583	1.367
Metro Urbanization	4.579	0.001	4.479	4.605	-1.638
Metro Literacy	4.412	0.001	4.367	4.459	0.064

Note: *Per capita GDP* is the logarithm of district per capita GDP in RS. Millions in 2001; *Distance* is the logarithm of the distance by road to the closest of the seven largest urban agglomerations as listed in Table 3; *Literacy* is the logarithm of the total literacy rate per hundred people; *Electricity* is the logarithm of the percentage of households with an electricity connection ; *Commercial Banks* is the logarithm of the number of commercial banks per thousand people; *Urbanization* is the logarithm of the percentage of urban households; *Irrigated land* is the logarithm of the net irrigated land area per million people; *Pucca Road* is the logarithm of the percentage of households connected by "Pucca Roads" ; *Metro Electricity* is the logarithm of the percentage of households with an electricity connection in closest metropolitan district; *Metro Urbanization* is the logarithm of the percentage of urban households in the closest metropolitan district; and *Metro Literacy* is the logarithm of the total literacy rate per hundred people in the closest metropolitan district.

Table 5a: Conditional Convergence and Distance

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Per capita GDP	0.0018	-0.0082**	-0.0088**	-0.0106***	-0.0124***	-0.0128***	-0.0129***	-0.0112***
	(0.0030)	(0.0037)	(0.0037)	(0.0038)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
Distance	-0.0055***	-0.0047***	-0.0047***	-0.0039***	-0.0042***	-0.0040**	-0.0040**	
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0016)	(0.0016)	
Urbanization		0.0103***	0.0099***	0.0070***	0.0068***	0.0065***	0.0067***	0.0065***
		(0.0021)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Literacy			0.0049	-0.0056	-0.0084	-0.0077	-0.0081	-0.0078
			(0.0071)	(0.0083)	(0.0085)	(0.0086)	(0.0087)	(0.0085)
Electricity				0.0125***	0.0122***	0.0127***	0.0123***	0.0138***
				(0.0043)	(0.0043)	(0.0044)	(0.0045)	(0.0044)
Commercial Banks					0.0064	0.0054	0.0053	0.0033
					(0.0048)	(0.0049)	(0.0049)	(0.0048)
Irrigated Land						-0.0035***	-0.0036***	-0.0039***
						(0.0013)	(0.0013)	(0.0012)
Pucca Road							0.0020	
							(0.0024)	
Metro Electricity								
Metro Literacy								
Metro Urbanization								
Gravity Parameter	3.0196	-0.5696***	-0.5312***	-0.3678***	-0.3399***	-0.3106***	-0.3087***	
	(1.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
Constant	0.0975***	0.1521***	0.1379***	0.1533***	0.2480***	0.2024**	0.1956**	0.1389*
	(0.0358)	(0.0385)	(0.0465)	(0.0466)	(0.0887)	(0.0868)	(0.0861)	(0.0827)
BP Test	0.47	2.40	2.44	1.58	2.06	4.26**	6.00**	4.93**
	(0.4921)	(0.1210)	(0.1185)	(0.2089)	(0.1515)	(0.0391)	(0.0143)	(0.0264)
F Test	24.23***	32.62***	31.60***	33.86***	43.22***	19.42***	10.30***	24.11***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	566	556	556	556	556	546	544	550
R-squared	0.3354	0.3841	0.3847	0.3974	0.3995	0.4039	0.4031	0.3989

Note 1: *, **, *** denotes 10, 5 and 1 percent levels of significance respectively.
Note 2: Robust (White) standard errors are used.
Note 3: F-tests are joint tests for State dummy variables.

Table 5b: Conditional Convergence and Distance including Metro variables

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Per capita GDP	0.0014	-0.0085**	-0.0094**	-0.0111***	-0.0125***	-0.0130***	-0.0130***	-0.0116***
	(0.0030)	(0.0037)	(0.0037)	(0.0035)	(0.0043)	(0.0042)	(0.0042)	(0.0042)
Distance	-0.0049***	-0.0039***	-0.0039**	-0.0032*	-0.0035**	-0.0033**	-0.0032**	
	(0.0016)	(0.0015)	(0.0015)	(0.0017)	(0.0016)	(0.0016)	(0.0016)	
Urbanization		0.0103***	0.0098***	0.0070***	0.0068***	0.0066***	0.0067***	0.0066***
		(0.0021)	(0.0021)	(0.0022)	(0.0022)	(0.0021)	(0.0022)	(0.0021)
Literacy			0.0068	-0.0041	-0.0062	-0.0053	-0.0057	-0.0054
			(0.0075)	(0.0076)	(0.0091)	(0.0093)	(0.0093)	(0.0092)
Electricity (%)				0.0124***	0.0122***	0.0125***	0.0121***	0.0133***
				(0.0038)	(0.0044)	(0.0045)	(0.0045)	(0.0044)
Commercial Banks					0.0047	0.0039	0.0038	0.0020
					(0.0049)	(0.0049)	(0.0049)	(0.0048)
Irrigated Land						-0.0036***	-0.0036***	-0.0039***
						(0.0012)	(0.0012)	(0.0012)
Pucca Road							0.0019	
							(0.0025)	
Metro Electricity	0.2349	0.2586	0.2914*	0.2508	0.2433	0.2561	0.2537	0.2535
	(0.1602)	(0.1600)	(0.1678)	(0.1588)	(0.1692)	(0.1735)	(0.1731)	(0.1746)
Metro Literacy	0.0966*	0.1084**	0.1082	0.1017*	0.0970*	0.1018**	0.0998*	0.1261**
	(0.0500)	(0.0500)	(0.0501)	(0.0530)	(0.0502)	(0.0515)	(0.0517)	(0.0503)
Metro Urbanization	-0.0837*	-0.0704	-0.0723	-0.0811*	-0.0764	-0.0720	-0.0730	-0.0740
	(0.0485)	(0.0473)	(0.0475)	(0.0451)	(0.0473)	(0.0485)	(0.0484)	(0.0492)
Gravity Parameter	3.4018	-0.4596***	-0.4200***	-0.2909***	-0.2818***	-0.2515***	-0.2509***	
	(1.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
Constant	-1.0146	-1.1826	-1.3408*	-1.0702	-0.9662	-1.1100	-1.0914	-1.2492
	(0.7672)	(0.7152)	(0.8147)	(0.8310)	(0.8293)	(0.8432)	(0.8416)	(0.8447)
BP Test	0.62	2.58	2.79	1.63	2.11	4.24**	6.07**	4.46**
	(0.4297)	(0.1081)	(0.0951)	(0.2020)	(0.1466)	(0.0394)	(0.0137)	(0.0348)
F Test	58.98***	27.04***	24.98***	28.61***	28.45***	14.59***	10.29***	16.46***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	566	556	556	556	556	546	544	550
R-squared	0.3436	0.3928	0.3939	0.4062	0.4073	0.4118	0.4109	0.4094

Note 1: *, **, *** denotes 10, 5 and 1 percent levels of significance respectively.
Note 2: Robust (White) standard errors are used.
Note 3: F-tests are joint tests for State dummy variables.

Table 6a: Stability Test for models excluding metro variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
North East	0.18	0.35	0.57	0.57	0.61	0.34	0.35	0.45
	(0.84)	(0.79)	(0.68)	(0.73)	(0.72)	(0.93)	(0.95)	(0.85)
Maharastra	0.40	1.20	1.26	1.22	0.95	1.07	1.20	1.33
	(0.67)	(0.31)	(0.28)	(0.30)	(0.46)	(0.38)	(0.30)	(0.24)
Bihar	0.93	0.42	0.35	1.67	2.83**	2.39**	2.24**	2.84***
	(0.40)	(0.74)	(0.84)	(0.14)	(0.01)	(0.02)	(0.02)	(0.01)
North East and Bihar	1.07	0.72	0.71	0.97	0.97	0.70	0.72	0.82
	(0.35)	(0.54)	(0.59)	(0.44)	(0.45)	(0.67)	(0.68)	()0.55
Maharastra and Bihar	0.81	1.23	1.04	1.82	1.62	1.49	1.49	1.23
	(0.44)	(0.30)	(0.39)	(0.11)	(0.14)	(0.17)	(0.16)	(0.29)
Karnataka	0.81	1.44	1.52	2.33**	1.80*	1.48	1.50	1.07
	(0.44)	(0.23)	(0.20)	(0.04)	(0.09)	(0.17)	(0.16)	(0.38)

Note 1 P-values are given in the parenthesis.

Note 2: F-tests are joint tests for sate dummy variables.

Table 6b: Stability Test for models including metro variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
North East	0.51	0.75	0.93	0.88	0.81	0.50	0.51	0.50
	(0.73)	(0.56)	(0.46)	(0.51)	(0.58)	(0.86)	(0.87)	(0.83)
Maharastra	0.15	0.64	0.71	0.58	0.48	0.66	0.74	0.72
	(0.93)	(0.63)	(0.62)	(0.75)	(0.85)	(0.73)	(0.68)	(0.65)
Bihar	0.94	0.43	0.38	1.69	2.84***	2.41**	2.24**	2.85***
	(0.39)	(0.73)	(0.82)	(0.13)	(0.01)	(0.02)	(0.02)	(0.01)
North East and Bihar	0.69	0.70	0.75	0.97	0.94	0.70	0.71	0.86
	(0.60)	(0.59)	(0.59)	(0.44)	(0.48)	(0.69)	(0.70)	(0.54)
Maharastra and Bihar	0.37	0.65	0.60	1.25	1.19	1.15	1.17	1.04
	(0.77)	(0.63)	(0.70)	(0.28)	(0.31)	(0.33)	(0.31)	(0.40)
Karnataka	1.54	1.60	1.94*	2.08**	1.74*	1.53	1.53	1.41
	(0.19)	(0.16)	(0.07)	(0.04)	(0.09)	(0.13)	(0.12)	(0.19)

Note 1 P-values are given in the parenthesis.

Note 2: F-tests are joint tests for sate dummy variables.

Table 7a: Results for Endogeneity Test for models excluding metro variables

	(2)	(3)	(4)	(5)	(6)	(7)	(8)
District Literacy	NA	22.70	13.21	12.36	9.21	9.90	9.62
		(0.8599)	(0.9986)	(0.9996)	(1.000)	(1.000)	(1.000)
District Pucca Road	NA	NA	NA	NA	NA	4.82	NA
						(1.000)	
District Urbanization	58.98***	55.37***	33.77	33.65	22.26	23.14	22.78
	(0.0012)	(0.0046)	(0.3818)	(0.4356)	(0.9216)	(0.9204)	(0.8850)

Note 1: *, **, *** denotes 10, 5 and 1 percent levels of significance respectively.

Note 2: All tests follow χ^2 with appropriate degrees of freedom equal to the number of model parameters.

Note 3: Endogeneity tests are performed by comparing OLS and 2SLS parameter estimates. This comparison is done in SUR framework. The Hausman test is not appropriate as data has heteroscedasticity.

Table 7b: Results for Endogeneity Test for models including metro variables

	(2)	(3)	(4)	(5)	(6)	(7)	(8)
District Literacy	NA	18.79	9.95	9.38	6.22	6.45	6.07
		(0.9839)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
District Pucca Road	NA	NA	5.61***	NA	5.59***	3.12	
			(1.000)		(1.000)	(1.000)	
District Urbanization	55.38***	50.07**	30.64	30.73	20.42	21.32	20.70
	(0.0087)	(0.0372)	(0.6787)	(0.7174)	(0.9829)	(0.9817)	(0.9737)

Note 1: *, **, *** denotes 10, 5 and 1 percent levels of significance respectively.

Note 2: All tests follow χ^2 with appropriate degrees of freedom equal to the number of model parameters.

Note 3: Endogeneity tests are performed by comparing OLS and 2SLS parameter estimates. This comparison is done in SUR framework. The Hausman test is not appropriate as data has heteroscedasticity.

Appendix: Variance Decomposition

This appendix briefly describes our variance decomposition. Let y_{ij} be the underlying variable (say, per capita logged income) of j^{th} district in i^{th} state, $j = 1, 2 \ldots, n_i, \quad i = 1, 2 \ldots K$. Let $N = \sum_{i=1}^{n_i}$, the total number of observations. Define $\bar{\bar{y}} = \frac{1}{N} \sum_{i=1}^{K} \sum_{j=1}^{n_i} y_{ij}$, the Grand mean. Define $\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, i = 1, 2 \ldots K$, the within mean. We define following three quantities...

Total sum of square (TSS)= $\sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \bar{\bar{y}})^2$.

Within Sum of square (WSS)= $\sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$.

Between Sum of Square (BSS)= $\sum_{i=1}^{K} n_i (\bar{y}_i - \bar{\bar{y}})^2$.

Then

$$TSS = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \bar{\bar{y}})^2 = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i + \bar{y}_i - \bar{\bar{y}})^2 = WSS + BSS.$$

Finally dividing each term by N gives the total, between and within-state variances, $\nu^T = TSS/N$, $\nu^W = WSS/N$ and $\nu^B = BSS/N$. Hence $\nu^T = \nu^W + \nu^B$.

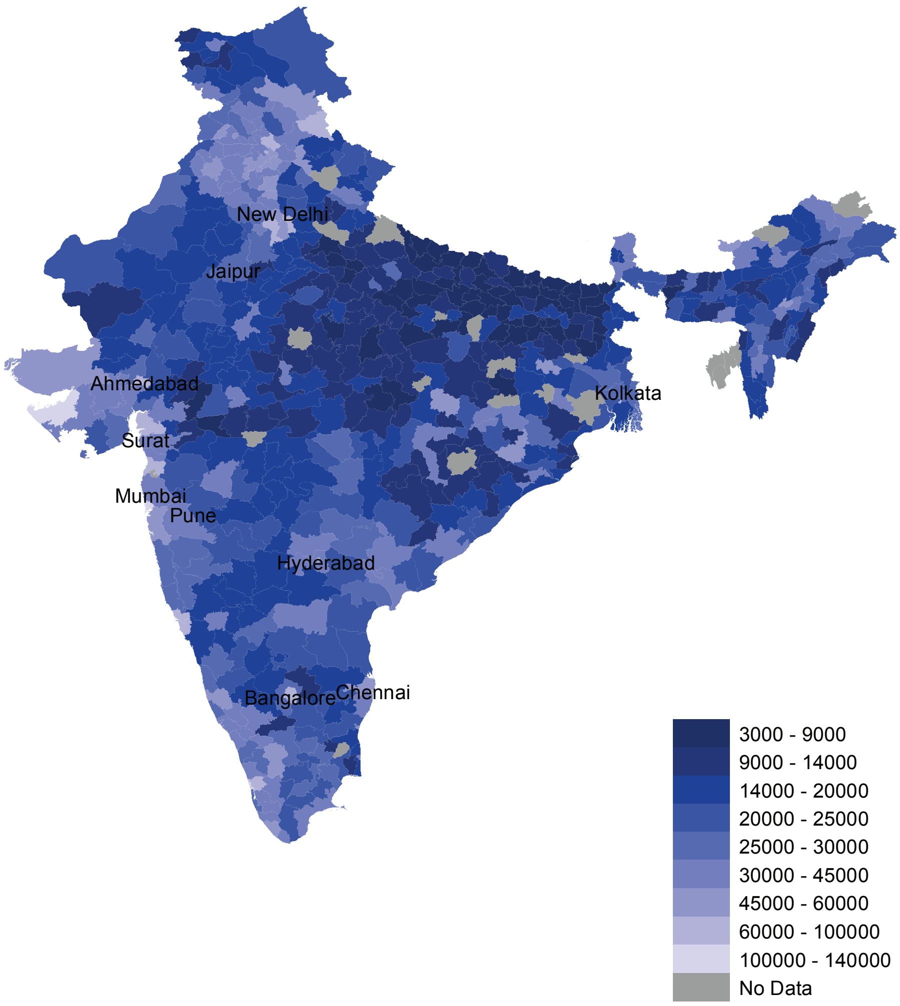

Figure 1: Per Capita Income by District

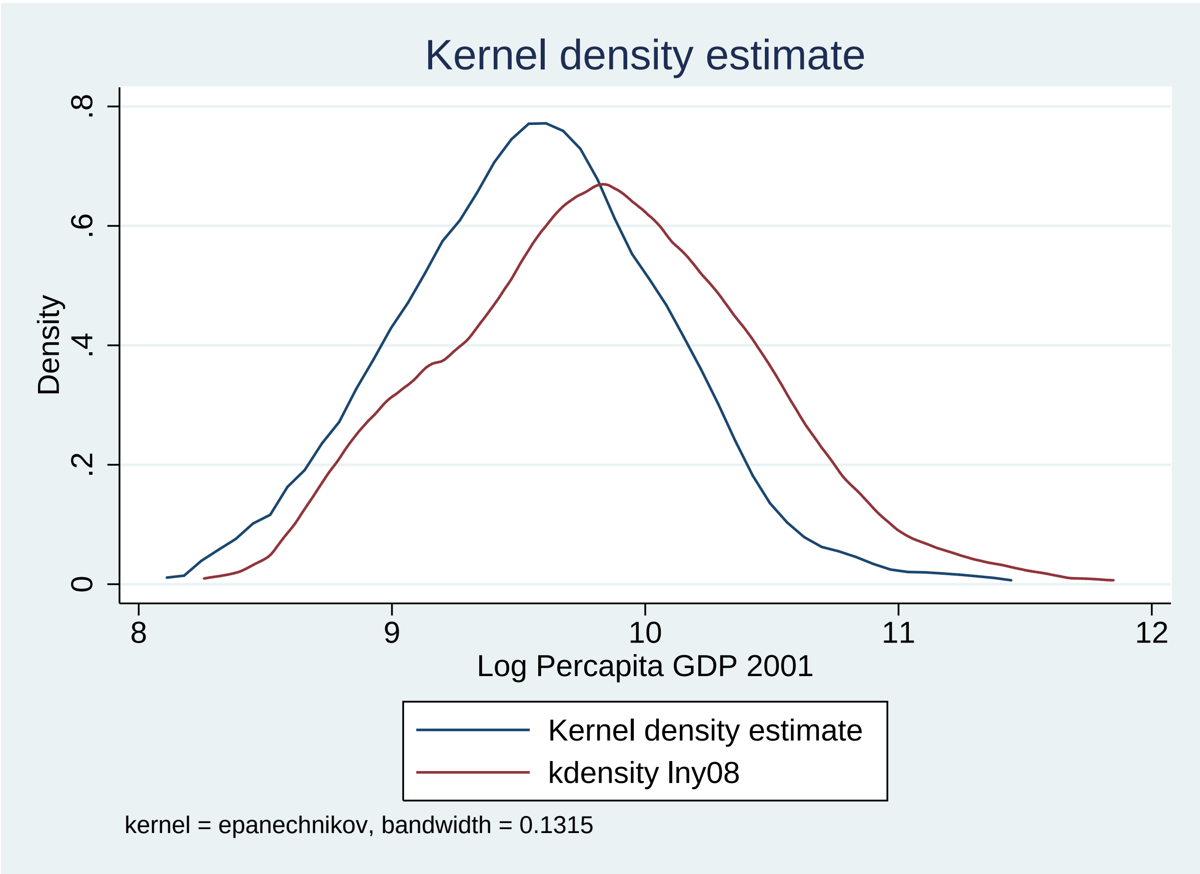

Figure 2: Probability Density Function for Indian District Incomes

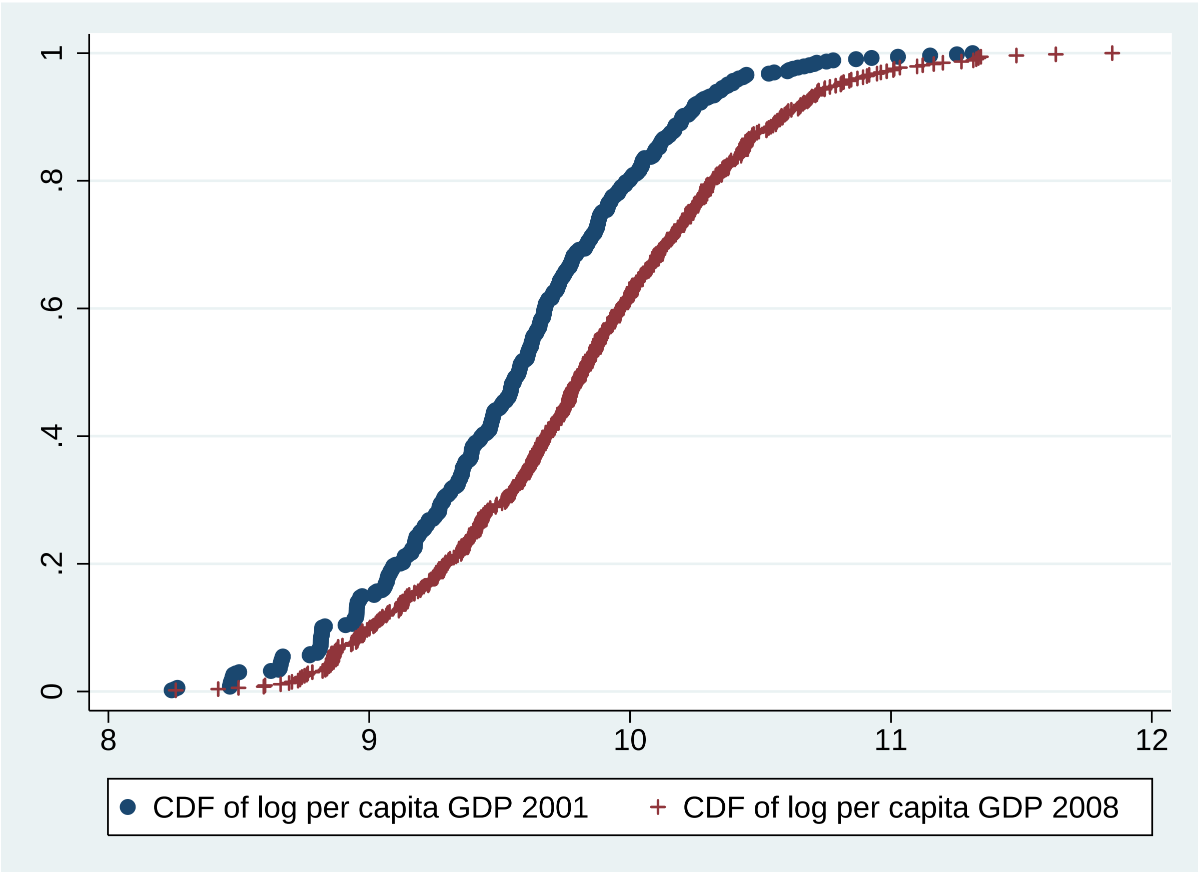

Figure 3: Cumulative Distribution Function for Indian District Incomes

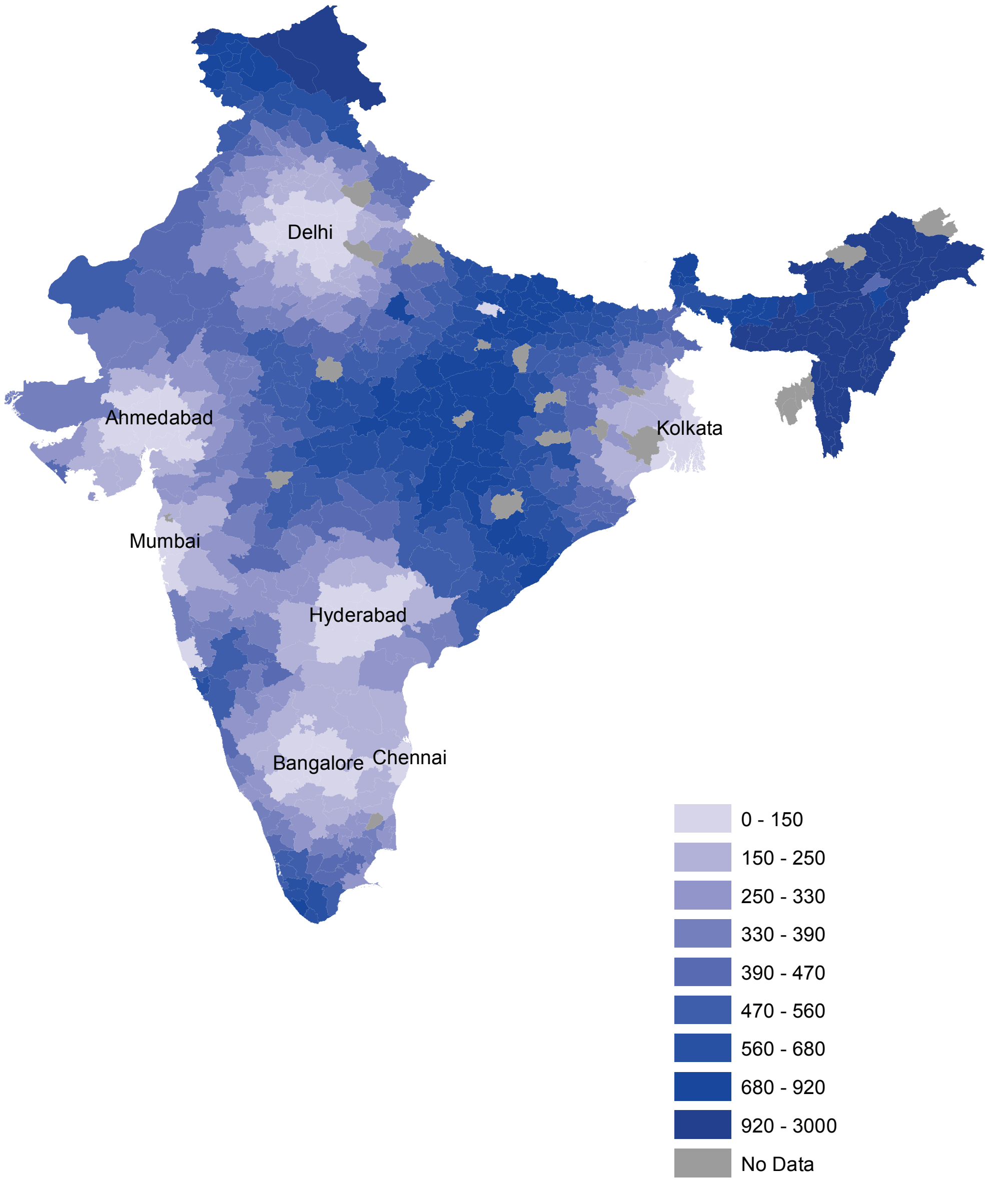

Figure 4: Minimum Distance to Seven Largest Metropolitan Centers

LATEST ICRIER'S WORKING PAPERS

NO.	TITLE	AUTHOR	YEAR
267	NORMALIZING INDIA PAKISTAN TRADE	NISHA TANEJA MISHITA MEHRA PRITHVIJIT MUKHERJEE SAMRIDHI BIMAL ISHA DAYAL	SEPTEMBER 2013
266	RECESSION AND CHILD LABOR: A THEORETICAL ANALYSIS	SAHANA ROY CHOWDHURY	APRIL 2013
265	IMPACT OF MACRO-ECONOMIC ENVIRONMENT ON DIVERSIFICATION-PERFORMANCE RELATIONSHIP: A CROSS COUNTRY STUDY OF INDIA AND JAPAN	SAPTARSHI PURKAYASTHA	MARCH 2013
264	FACTOR INCOME TAXATION, GROWTH, AND INVESTMENT SPECIFIC TECHNOLOGICAL CHANGE	MONISANKAR BISHNU, CHETAN GHATE AND PAWAN GOPALAKRISHNAN	MARCH 2013
263	INDIA'S ROLE IN FACILITATING TRADE UNDER SAFTA	NISHA TANEJA SHRAVANI PRAKASH PALLAVI KALITA	JANUARY 2013
263	INDIA'S ROLE IN FACILITATING TRADE UNDER SAFTA	NISHA TANEJA SHRAVANI PRAKASH PALLAVI KALITA	JANUARY 2013
262	SECTORAL INFRASTRUCTURE INVESTMENT IN AN UNBALANCED GROWING ECONOMY: THE CASE OF INDIA	CHETAN GHATE GERHARD GLOMM JIALU LIU	NOVEMBER 2012
261	WHY WAS THE PARTICIPATION OF INDIAN STATES IN THE GROWTH TURNAROUND SO PATCHY? SOME EVIDENCE BASED ON ROBUSTNESS ANALYSIS	CHETAN GHATE AND STEPHEN WRIGHT	OCTOBER 2012
260	HAS INDIA EMERGED? BUSINESS CYCLE STYLIZED FACTS FROM A TRANSITIONING ECONOMY	CHETAN GHATE RADHIKA PANDEY ILA PATNAIK	JUNE 2012

About ICRIER

Established in August 1981, ICRIER is an autonomous, policy-oriented, not-for-profit, economic policy think tank. ICRIER's main focus is to enhance the knowledge content of policy making by undertaking analytical research that is targeted at informing India's policy makers and also at improving the interface with the global economy. ICRIER's office is located in the institutional complex of India Habitat Centre, New Delhi.

ICRIER's Board of Governors includes leading academicians, policymakers, and representatives from the private sector. Dr. Isher Ahluwalia is ICRIER's chairperson. Dr. Rajat Kathuria is Director and Chief Executive.

ICRIER conducts thematic research in the following seven thrust areas:

- Macro-economic Management in an Open Economy
- Trade, Openness, Restructuring and Competitiveness
- Financial Sector Liberalisation and Regulation
- WTO-related Issues
- Regional Economic Co-operation with Focus on South Asia
- Strategic Aspects of India's International Economic Relations
- Environment and Climate Change

To effectively disseminate research findings, ICRIER organises workshops, seminars and conferences to bring together academicians, policymakers, representatives from industry and media to create a more informed understanding on issues of major policy interest. ICRIER routinely invites distinguished scholars and policymakers from around the world to deliver public lectures and give seminars on economic themes of interest to contemporary India.