ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

RICH STATES, POOR STATES: CONVERGENCE AND POLARISATION IN INDIA

Sanghamitra Bandyopadhyay

Number 266

June 2006

Manor Road Building, Oxford OX1 3UQ

Rich States, Poor States: Convergence and Polarisation in India[1]

Sanghamitra Bandyopadhyay
Department of Economics and Oriel College, Oxford University
and STICERD, London School of Economics

11th April 2006

[1]I would like to thank Henry Overman, Danny Quah and Diana Weinhold for guidance, and participants at seminars at the London School of Economics, Royal Holloway College, University of Houston and University of Oxford, and Chris Adam, Frank Cowell, Peter Davis, Marcel Fafchamps and Tony Venables for useful comments. Funding from the Economic and Social Research Council, UK, is gratefully acknowledged. All errors are mine. **Correspondence**: sanghamitra.bandyopadhyay@economics.ox.ac.uk

Abstract

The distribution dynamics of incomes across Indian states are examined using the entire income distribution rather than using standard regression approaches.The period 1965 to 1997 exhibits twin-peaked dynamics: there are two income convergence clubs at 50% and 125% of the national average income. Disparities across the states declined over the sixties and then increased. The observed polarisation is explained by the disparate distribution of infrastructure, in particular, that of education, irrigation and literacy in the formation of the lower convergence club. Parametric analysis establishes irrigation, education, roads, industrial power consumption and bank deposits as infrastructure components explaining cross-state variation in growth.

Keywords: Convergence clubs, distribution dynamics, education, infrastructure, panel data, India

JEL Classification: C23, E62, O23

Word Count: 9,079

1 Introduction

This paper documents the dynamics of growth and incomes across Indian states over three decades (1965-1997). The evolution of the state-level income distribution is examined rather than investigating whether one obtains beta or sigma convergence. This is necessary not only to understand the empirics of catch-up more accurately but also to identify the nature of the underlying distributional dynamics: whether there are long-run cohesive tendencies, polarization, stratification or the emergence of convergence clubs. Some explanations are then provided - we find some infrastructural indicators to significantly explain the evolution of the income distribution and the formation of convergence clubs.

Some simple statistics reveal the wide disparities in growth across Indian states - Punjab's income has been at least twice that of Bihar's, Orissa's and Rajasthan's since 1965. Some states have doubled their incomes (real GDP per capita) over the period of the mid sixties to the 1990s, while the poorest states lie well below the national average income. The analysis in the paper reveals twin-peaked dynamics: that while over the late 1960s there were some tendencies of cohesion, over the 1970s through to the 1990s incomes have persistently polarised into two convergence clubs. Finally, we will identify some auxiliary factors which explain the observed dynamics. We focus, in particular, on the role of a set of economic and social infrastructural indicators in explaining the observed polarisation. The results suggest that some infrastructural indicators, namely that of literacy, extent of irrigated land and transport infrastructure robustly explain the formation of the lower convergence club.

Of course, a large literature already studies unequal growth across Indian states. Those based on the popular cross section regression approach, of Bajpai and Sachs (1996), Cashin and Sahay (1996), Nagaraj and Venganzones (1997), Aiyar (2000) and Trivedi (2003) emphasise such diverging distributional characteristics, but none about intradistributional mobility. Convergence as an empirical concept, as defined by Solow (1956) is understood as a single economy approaching its theoretically derived steady state growth path; so standard empirical analyses, thus, only study the behaviour of the single (representative) economy. While such an empirical methodology can accurately uncover tendencies of divergence, it does not uncover the empirical regularities of the distributional patterns (of polarisation or stratification) that we wish to expose. Similarly, time series approaches, (Carlino and Mills (1993), Loewy and Papell (1994)) which estimate the univariate dynamics of income also remains incomplete in describing the dynamics of the entire cross section.

Recent studies are coming to recognise the non-linear nature of the effects of various explanatory factors of such income dynamics - Kalaitzidakis and Savvides (2001) for instance, highlights the non-linear nature of education on economic growth. Standard empirical tools of panel or cross section regression are not fashioned to explain the formation of clubs at different parts of the income distribution. Theoretical contributions of (Bernaud and Durlauf (1994), Durlauf and Johnson (1994), Esteban and Ray (1994), Ben-David (1994), De Long (1988), Galor and Zeira (1993)) also allow for explicit patterns of cross-economy interaction, whereby economies cluster together into groups to endogenously emerge. They recognise that economies do not evolve in isolation, but in clubs and groups, and such distributional characteristics remain unexposed under standard empirical techniques for studying convergence.

Bianchi (1995), Jones (1997), Desdoigts (1994) and Fiaschi (2003) study the evolution of the entire distribution over time using various non-parametric methods, while Quah (1997) formalizes empirical models for estimating the distribution dynamics - this is the approach adopted. Markov chains are used to approximate and estimate the laws of motion of the evolving distribution. The intra-distribution dynamics information is encoded in a transition probability matrix, and the ergodic distribution associated with this matrix describes the long term behaviour of the income distribution. It encompasses both time series and cross section properties of the data simultaneously and presents itself as an ideal approach for large data sets. Moreover, this method can be extended to identify factors governing the formation of these convergence clubs. Identification of some explanatory factors will comprise a major part of the paper and contributes to a growing body of empirical literature on the identification of non-linear effects of various factors, as in our case infrastructure, on economic growth.

The rest of the paper is organised as follows. Section 2 discusses some simple dynamics and introduces the distribution dynamics approach. Section 3 describes the observed distribution dynamics and polarisation. Section 4 discusses and presents various explanatory factors which explain the distribution dynamics. Section 5 presents results of the various conditioning schemes using standard panel regressions to contrast the results obtained with distribution dynamics. Section 6 concludes.

2 Some simple dynamics

Let us take a cursory look at some basic statistics. GDP per capita and price data used for this paper has been obtained from Özler and Ravallion (1996). GDP per capita data for 1989 to 1998 has also been obtained from the World Bank, compiled as a separate dataset, but from same Government of India sources.

The richest state, Punjab, already had a per capita income of 270 (in 1990 dollars for international comparability) in 1965, which increased to 370, increasing by a factor of 34% by 1988, and by another 21% by 1997. Another group of rich states,Gujarat's and Maharashtra's per capita income had increased from 183 and 196 (in 1990 dollars) to 233 and 303 by a factor of 20% and 27 %, and by another staggering 40% and 51% by 1997, respectively. By comparison, the Indian average per capita GDP (in 1990 dollars) was 153 in 1965 and 195 in 1988 (increasing by 27 %), and increasing by another 33 per cent by 1997 . Hence, Punjab was already almost twice as rich as the Indian average in 1965 and remained so at the end of the period. Maharashtra, Gujarat and Haryana's income per capita have also maintained a per capita income of almost twice the Indian average all throughout the period. Averaging, states of Punjab, Haryana, Gujarat, Maharashtra were at 123%, in 1965 and over 152%, in 1988 of the Indian average and grew another 36% as a group over 1988 to 1997.

The poorest regions are also evident - Bihar, Orissa in the east, Rajasthan in the west, and Uttar Pradesh in the north have consistently been lying around the lowest per capita GDPs. Bihar, Orissa and Uttar Pradesh and Rajasthan have been at 85% in 1965 and 80% in 1988 of the Indian average, and by 1997 grew only another 19% as a group. Bihar and Orissa had per capita GDPs of 122 and 121 in 1965 and 122 and 145 in 1988 (in 1990 dollars). Thus over the entire period of study, the income of the richer states has been almost three times that of the poor. Interestingly, while the growth rates of Madhya Pradesh, Assam, Andhra Pradesh, Uttar Pradesh, Orissa, and Bihar, the six poorest states, were all significantly below the national growth rate, they account for more than half of the Indian population.

However, not all that were rich remained rich, and those poor remained poor. West Bengal, notably, with a GDP per capita of 196 in 1965 and 205 (in 1990 dollars) in 1988 fell steeply in its ranking from second to eighth by 1988. Thus, West Bengal teamed with Punjab, Haryana and Maharashtra as a high growth state in the 1960s, but experienced dismal growth over the following years. Again, while the surge of growth in the 1980s benefited the four richest states, it also pushed up Karnataka and Tamil Nadu, whose 1988 per capita income had increased by 21% and 36% over 1980-88, and a further

45 % for each, by 1997.

Analysing the same details reveals that over 1965 to 1997 the standard deviation (SD) of per capita income has increased by 192%, while the interquartile range (IQR) has increased by 137%. A significant increase in spread manifests clearly - the SD is about double that of the IQR. This, however, has an interesting implication. With the IQR accounting for the middle 50% of the distribution, the SD exceeding it (by such a large amount) can only be attributed to some high performers outperforming the rest of the intermediate and poor states.

These back-of-the-envelope calculations reveal the dynamic spatial patterns of regional growth in the Indian states. It reveals both persistence and mobility. While some rich have remained rich, and the poor persistently poor, there have been some signs of mobility - instances of high performers who have declined in their performance over the period - West Bengal, others who have picked up over the period, for example, Karnataka. Thus, apart from those consistent performers, there is plenty of evidence of relative successes and failures all across India.

2.1 Still More Dynamics

To reveal the empirical regularities of the intra-distributional dynamics over the given period of time, we track the evolution of the entire income distribution itself over time. Markov chains are used to approximate and estimate the laws of motion of the evolving distribution. The intra-distribution dynamics information is encoded in a transition probability matrix, and the ergodic (or long run) distribution associated with this matrix describes the long term behaviour of the income distribution.

The empirical model we will use will analyse the evolution of the income distribution; documenting the intra-distributional dynamics, club formation, and the long run tendencies. The distribution dynamics approach (due to Quah 1997) is based on treating a single income distribution as a random element in a field of income distributions, called the random field. The density function of the income distribution is non-parametrically estimated at each point in time and is then observed how it evolves over time. The primary tool used to track the evolution of the income distribution is the transition probability matrix, which will record the probabilities of persistence and mobility across the income distribution.

2.2 Empirical Models of Intra-distribution Churning

Both stochastic kernels and transition matrices provide an estimate of intra-distribution mobility taking place. In both cases, it is assumed that an economy (in our case, a state) over a given time period (say, one year or five years) either remains in the same position, or changes its position in the income distribution. Such a change in position of an economy in the income distribution is called a transition. Our task is to observe how many such transitions take place in the given time period.

First, what needs to be identified is the position of the economy in the income distribution in the starting period. This is done by dividing the income distribution into "income states". Income states are a range of income levels, say between a fifth and a half of the weighted average of the country. Then we observe how many of the economies which are in an income state say, (0.2, 0.5) in the initial period land up in that very state, or elsewhere in the next time period. If they do end up in another income state, (for example, in the income range of a half to three quarters of the weighted average income) there is said to be mobility. If they end up in the same, there is persistence. We will be interested in the former possibility i.e. of intra-distribution mobility.

In our exercise on India, we have measured these transitions and the results are tabulated in Tables 1a-d as transition probability matrices.

Interpreting the transition matrix is as follows (see Tables 1a-d): First, we discretise the space of possible values of income, in r states. For instance, we define the state i (0.2 , 0.5) as one which has regions with an income lying between 0.2 and 0.5 times the average income of the country. The probabilities obtained, give us the percentages of economies (in our case, Indian states) which given a starting state, have moved on to a different state. So, our row probabilities all add up to 1. Of these, the diagonal of the transition probability matrix is of interest to us. A diagonal with high values indicates high probabilities of persistence - a high likelihood of remaining in a particular state when one starts there. Thus, the smaller the diagonal, the greater intra-distribution mobility there exists.

The transition probability matrix also allows us to take a long run view of the evolution of the income distribution. This is tabulated in the row called the "Ergodic Distribution".

There is, however, a drawback in this measure as the selection of income states is arbitrary - different sets of discretisations may lead to different results. The stochastic kernel improves on the transition probability matrix by eschewing arbitrary discretisations, allowing the space of income values to be a continuum of states[1]. This means that we no longer have a grid of fixed income states, like (0.2 0.5), (0.5 0.75) etc. but allow the states to be all possible intervals of income. By this we remove the arbitrariness in the discretisation of the states. We now have an infinite number of rows and columns replacing the transition probability matrix.

Interpreting the stochastic kernels is as follows (see Figures 2a-d). Any slice running parallel to the horizontal axis (i.e. t + k axis) describes a probability density function describing the transitions from one part of the income distribution to another over k periods. The location of the probability mass will provide us information about the distribution dynamics, and thus about any tendencies of convergence. Concentration of the probability mass along the positive slope indicates persistence in the economies' relative position and therefore low mobility. The opposite, i.e. concentration along the negative slope, would imply overtaking of the economies in their rankings. Concentration of the probability mass parallel to the t + k axis indicates that the probability of being in any state at period t + k is independent of their position in period t i.e. evidence for low persistence. Finally, convergence is indicated when the probability mass runs parallel to the t axis.

3 The Distribution Dynamics.

We will now take a look at the distribution dynamics of incomes across Indian states over 1965 to 1997. Figures 1a to 1d present the stochastic kernels for per capita income (relative to national average) of 1-year transitions for four sub-periods 1965-70, 1971-1980, 1981-88, and 1990-97. Data on real per capita income has been obtained from Özler and Ravallion (1996)and the World Bank.

Observation of the stochastic kernels and the contour plots reveal that the later years provide increasing evidence of persistence and low probabilities of changing their relative position. Over the periods 1965-70, 1971-80, 1981-88, 1990-97 we observe in Fig. 1a-d the probability mass lengthening and shifting totally in line with the positive diagonal, the two peaks still at the two ends of the mass. The cluster of states at the two peaks consist of some low income economies at around 50 per cent of the all India average and another at 125 per cent of the average. Thus, though an overall view of the entire sample period 1965-97 shows some signs of cohesion, the sub-sample periods,

[1] Such refinement goes beyond the generalisation as well. It is well known that discretisation may well remove the Markov property from an otherwise well behaved Markov process, Chung (1960)

particularly during the later years, have shown cohesive forces substantially dissipating in influence. The result is that of the rich states forging ahead, with the poor making little progress and a dispersing middle income group.

For robustness, we estimate transition matrices and stochastic kernels over the different sub periods, and over longer (5 and 10 years) periods. The results obtained (not presented for brevity) suggest the same results as those above: the pressing facts that are revealed are that of convergence over the late 1960s, with increasing divergence over the 1970s, 1980s and 1990s.

The long run view of whether the economies will converge over the long run is addressed by estimating the transition probability matrices. The results are tabulated in Tables 1a-d. Interpretation of the tables is as follows. Each of the defined states for each table is different, such that each distribution is uniform at the beginning year of the sample[2]. The first column of the table accounts for the number of transitions over the time period beginning at each state. The following columns present the calculated probabilities of transition from one specified state to another. Like the stochastic kernel, a "heavy" main diagonal is bad news - i.e. indicating persistence.

Table 1a reports results for 1965-70 and they are quite similar to those obtained for the stochastic kernel - the values in the main diagonal are around 50%, which indicates that the probability that an economy remains in its own income state is around 50%. The off-diagonal values are those which are indicative of mobility, albeit little. Mobility is evident and obvious for the above average income group. The states with incomes in the first two states reveal some low income states which have forged ahead. We also have an estimator of the long run tendencies, named the ergodic distribution, accounted in the last row of the table. This gives us the long run tendency of an economy to land up in a given income range. The results suggest that over the long run, the probability that an economy lands up in the 4th state is the highest, a little over 40%. What is encouraging is that the lower income groups vanish in the ergodic distribution.

Following Tables 1b to 1d give us estimates of the transition matrix for the following sub-periods. The second period again reveals tendencies of both persistence and mobility, with tendencies of persistence in the lower income group and the high income groups. The probability that the first two income states and last two income states shift anywhere other than their own is zero. Though there are signs of persistence, there is evidence of some inter-state (income state) movement, again in the high income clusters. This

[2]It has been clarified by Bulli (2001) that a semi-Markovian model accentuates the results of persistence obtained using the current Markovian model with uniform initial distributions.

trend continues in the next two periods.

It is important to remember that as these estimates are based on time stationary transition matrices, they are not reliable for long time periods for economic structural changes. For further methods highlighting how more information may be obtained from such transition matrices, see Faischi et al (2003).

To summarise: using transitional matrices and stochastic kernels one finds evidence of persistence, and increasing divergence over the period 1970s to the 1990s, though some evidence of intra-distributional mobility over the late 1960s is revealed. The stochastic kernels in particular provide evidence of the formation of convergence clubs at two different points of the distribution - and their evolution over the three decades suggests that disparities have been widening and that forces of cohesion exist within the clubs revealed. With the sample size very small, robustness statistics (such as bootstrapped standard errors for the probabilities) are statistically of little relevance here. The evidence obtained is however in clear confirmation of the simple statistics discussed in section 2. Trivedi (2003) also highlights the formation of clubs with kernel estimates of the densities of the Indian state income distribution over 1960 to 1992. The stochastic kernels improve over these estimates by providing the pattern by which these clubs evolve.

4 Explanations of polarised economic growth across Indian states - identification of non-linearities

Differential growth and development across regions is very often attributed to different levels of infrastructure development. While the choice of a particular kind of infrastructure addressed in the paper can be considered as ad hoc in choice, the focus on the role of different kinds of infrastructure remains compelling, particularly for developing economies.

The growth-equipment investment nexus is found to be particularly strong in LDCs, and is attributed to the high returns to equipment investment in LDCs. Similar concerns regarding the importance of human and social infrastructure, embodied in the likes of literacy, and health and various organizational practises are widely discussed too. Theoretical studies of Stokey (1991), Galor and Tsiddon (1994) all associate human capital accumulation and economic growth positively. Empirical studies of Lee (1993)and Benhabib and Spiegel (1994) also evince that there exists a strong association between educational attainment and economic growth. The importance of financial infrastructure in aiding development and growth is also stressed. Empirical studies of King and Levine (1992), Benhabib and Spiegel (2000) highlight the positive association of finance and economic growth, and on India in particular, Burgess and Pande (2003) also find evidence of rural credit expansion in having increased Indian income per capita and reduced poverty over the 1960s to the 1990s.

Using both the distribution dynamics (non-parametric) method and standard panel regressions, we will address the role of different infrastructural indicators in having perpetuated the observed polarisation. The distribution dynamics method highlights the distributional effects of a particular infrastructure - does it affect the poor states the most, or is its effect global? Recent studies have already sought to identify such non-linear relationships using similar non-parametric methods, for example Kalaitzidakis and Savvides (2001)and Fiaschi (2003). Economic growth is increasingly viewed to be a non-linear process within new growth theory, though the majority of empirical analyses, by the use of cross section and panel regression analyses, have principally focussed on linear relationships. As a non-parametric methodology, the conditioning distribution dynamics thus serves to uncover any such non-linearities in the relationships, and thus the evolution of their relationship over time. Standard panel regressions accompanying the distribution dynamics, compare and ensure robustness of the results, and highlight the relationships which go unrevealed as non-linear in their effects.

4.1 Explaining the Distribution Dynamics.

How does one go about explaining the observed polarisation? In this section we will undertake some non-parametric tools (i.e that of distribution dynamics) to identify some explanatory factors. In the following section we will complement these results with some standard parametric methods.

The non-parametric tools used are those proposed by Quah (1997). While standard regression approaches explains average or representative behaviour, the distribution dynamics method explains the evolution of the entire distribution, hence exposing and explaining behaviour at different parts of the distribution. Thus, while standard methods compare E(Y) and E(Y|X), hence determining whether X explains Y, this approach maps the entire distribution of Y to Y|X. If there is no change in the distributions, conditioned and unconditioned, we then conclude that the auxiliary factor does not explain the polarisation (or any other observed distribution pattern). However, if it does explain the polarisation, the distribution will have changed, where all economies in the conditioned distribution have the same income. This will all be revealed in the two models which are used in this method, described in the following section.

How will all this be revealed in the stochastic kernels and transition probability matrices? Mappings obtained earlier to observe the distribution dynamics characterise transitions over time. It can further be shown (see Quah 1997) that just as stochastic kernels (and transition matrices) can provide information about how distributions evolve over time, they can also describe how a set of conditioning factors alter the mapping between any two distributions. Hence, to understand if a hypothesised set of factors explains a given distribution we can simply ask if the stochastic kernel transforming the unconditional one to the conditional one removes those same features.

One extreme situation would be where we find that the mapping from the unconditional to the conditional distribution would have the probability mass running parallel to the original axis at one. This would mean that all states irrespective of their own income would have their income conditioned by the auxiliary factor close to one. Since all incomes here are relative to the national average, this would mean that income, once conditioned, leads to "conditional convergence" - where all incomes converge to the national average. The conditioning factor would therefore be deemed as a factor explaining the observed polarisation. This, of course, is the desired outcome.

Another extreme would be where the stochastic kernel mapping the unconditional income distribution to that conditioned has its probability mass running along the diagonal. Unlike the previous case, this now implies the opposite possibility - each state, irrespective of its position in the initial distribution, has its income conditioned by the auxiliary factor unchanged. This renders the conditioning factor as one which does not explain the observed polarisation.

4.2 Conditioning on various infrastructural indicators

4.2.1 The data

The following infrastructural indicators[3] (panel data) are used for the analysis the analysis are the following. The states covered for the analysis are stated in the Appendix, and the period of study is 1977-1993. There are no missing observations.

Per capita electrical consumption (in kilowatt hours)

Per capita industrial consumption of electricity (in kilowatt hours)

Percentage of villages electrified.

[3]The infrastructure indicators' data set has been provided by the India team, Development Centre, OECD, Paris. The author gratefully acknowledges thanks to A. Varoudakis and M.Veganzones for kindly providing the data set.

Percentage of gross cropped area irrigated
Road length (in kms per 1,000 square kms)
Number of motor vehicles per 1,000 population.
Rail track length (in kms per 1,000 sq.kms)
Literacy rates (in percentage of the age group)
Primary school enrolment (age 6-11, in percentage of the age group)
Secondary school enrolment (age 11-17, in percentage of the age-group)
Infant mortality in percentage)
Number of bank offices per 1,000 population
Bank deposits as a percentage of the SDP
Bank credit as a percentage of the SDP

To generalise our results on infrastructure, we construct a single index accounting for the each of the state's infrastructural base. We use factor analysis to obtain the general index of infrastructure. This technique is a method of data reduction and attempts to describe the indicators as linear combinations of a small number of latent variables.We accept the first factor (f1, which we will call INFRA) to be the general index of infrastructure, which takes an eigenvalue of over 12 (out of a total of 17 indicators), results in Table 2[4].

The distribution dynamics of the index INFRA in Figure 2 sheds some interesting light on the change in its distribution. Though the upper half of the probability mass lies on the diagonal, the bottom half twists sharply anti-clockwise and runs parallel to the vertical line passing through 1. This implies that lower income group states have observed convergence in their levels of infrastructure. While this result does not shed any light on its role in explaining the observed polarisation of incomes across Indian states, it highlights how poorer states have had similar levels of infrastructure, an insight which will be useful later on.

We will now construct the conditioned distribution with various infrastructure variables. We confirm endogeneity of the infrastructural variables we are using for this study by Granger causality tests (results not reported for brevity). We cannot, thus, include the respective infrastructure variables as an exogenous variable in our growth equations but need to estimate the appropriate conditional distribution free from the feedback effects.

The conditioned distribution is obtained by regressing growth rates on a two-sided distributed lag of the time varying conditioning variables and then extracting the fitted residuals for subsequent analysis. This will result in a relevant conditioning distribution irrespective of the exogeneity of the right hand side variables. The method derives from that suggested by Sims

[4]See Bandyopadhyay (2004)

(1980), implemented in Quah (1996), where endogeneity (or the lack of it) is determined by regressing the endogenous variable on the past, current and future values of the exogenous variables, and observing whether the future values of the exogenous variables have significant zero co-efficients. If they are zero, then one can say that there exists no "feedback", or bi-directional causality. Needless to say, the residuals resulting from such an exercise would constitute the variation of the dependent variable unexplained by the set of exogenous variables, irrespective of endogeneity. We present the results for these two-sided regressions in Table 3.

What is observed in all projections is that infrastructure at lead 1 though lag 2 appears significant for predicting growth, but other leads and lags, not so consistently. Fit does not seem to improve with increasing lags (or leads). We seem to have a fairly stable set of co-efficients of the two-sided projections. The residuals of the second lead-lag projections are used as the conditioned distribution of growth on infrastructure, though results are unaltered by using residuals from other projections.

Fig 3a and the contour in Fig. 3b. suggest, the inter-state distribution of infrastructure have a role to play in explaining the polarisation of income across the states. One observes conditional convergence for the lower income convergence club.

To track the individual role of each infrastructure indicator, we now track the conditioning distribution dynamics of each indicator, of these results of the relevant variables are reported.

Fig. 4 presents the stochastic kernel mapping each state's income (relative to the national average) to that relative to the average income of states with the same level of education. We construct a composite index of education by factor analysis, results not presented, using three indicators of educational attainment - percentage of the population literate, primary school enrolment rates, secondary school enrolment rates. Once again, we run the two-sided lead-lag regressions to account for possible endogeneity, and extract the residuals to obtain the relevant conditioned distribution. The conditioning results obtained are encouraging - we find that for the lower income states the kernel twists anticlockwise, running fairly parallel to the "original" axis. Most of the upper half of the kernel runs along the diagonal. This implies that for lower income groups, and at the very upper end of the income distribution, education does explain the evolution of a state's SDP.

Interesting results are obtained when the conditioning is performed only with literacy (percentage of population literate) as the auxiliary factor. Fig. 5 reveals that the lower half of the stochastic kernel twists anticlockwise and runs parallel to the original axis. One thus obtains evidence of conditional convergence for the lower income club. Of the other infrastructure indicators, one obtains encouraging results for conditioning with percentage of irrigated land, as revealed in Figure 6. Here, one obtains evidence of conditional convergence in the lower tail of the kernel.

Could differing levels of state development expenditure[5] be responsible for differential development across the states? Here again, we again generate residuals from the lead-lag regressions of growth of GDP on state development expenditures, which serve to be the conditioned distribution. Figure 7 presents the stochastic kernel for the state development expenditure conditioning - the dominant features that characterise the kernel is that of the probability mass running mainly along the diagonal, indicating persistence and immobility for the most of the income distribution. A closer look, however, reveals that at higher income levels (those above the national average) and below 50% of the national average, the kernel twists anticlockwise. This implies that state domestic expenditure marginally affect the dynamics of the distribution at the higher and lower ends.

The relative insignificance of state development spending in our estimates does not necessarily mean that such spending is irrelevant to progress in reducing growth disparities, since other significant variables in the model may themselves be affected strongly by development spending. The impact of roads, education and infant mortality may well be reflecting in part the development spending on physical and social infrastructure.

We have thus uncovered some significant empirical correlations between a number of auxiliary factors and the observed income dynamics. The results obtained are distinctly different from those obtained in earlier studies in that it isolates "conditional convergence" at specific parts of the income distribution - results which would go uncovered in standard methods of investigating for conditional convergence using regression analyses. We now complement these results with those of standard parametric specifications in the following section both for robustness and to highlight the differences in the two methodologies' results.

5 Conditioning with parametric specifications

We now complement the earlier results with those derived from some standard parametric specifications to confirm the robustness of our results. Addressing both sets of results will enable us to identify the extent of the information that one can derive from..Focusing on the evolution of the distribution as a whole allowed us to observe the role of various auxiliary factors

[5] See Appendix for details for the composition of development expenditure.

at different levels of the distribution. Our earlier conditioning schemes explaining the distribution dynamics reveal that different conditioning criteria have mattered at different parts of the distribution. For example, conditioning schemes with infrastructure and state development expenditure, seem to explain polarisation at the two tails of the distribution, with no information on those of the middle-income states. Explaining differential behaviour at different levels of the income distribution is particularly important for policy purposes in targeting specific states with particular development strategies.

To complement our non-parametric results, we now propose some parametric specifications.

For each state, $i = 1, ..., N$ over dates $1, ..., T$ we estimate a growth regression given by

$$\ln Y_{it} - \ln Y_{it-1} = \alpha + \gamma X_{it} + \varepsilon_{it} \tag{1}$$

where the dependent variable is the growth rate of per capita income of state i in year t, α_i is a state- specific effect, X_i is a vector of regional characteristics, comprising of initial conditions and trends in exogenous time-dependent explanatory variables. The explanatory variables which are used in the analysis are as follows.

-share of agriculture in state domestic product

-rate of inflation measured as the change per year in the natural log of the (adjusted) CPIAL

-infrastructure (measured as INFRA, calculated earlier, incorporating both physical and social infrastructure. Individual infrastructural indicators as described earlier will also be used in separate regressions.

-real state development expenditure per capita.

We account for differences in production structure across states by introducing the share of agriculture in SDP as a control variable. We also control for inflation; the adverse and disparate impact of inflation on regional growth has been identified in past research (Bell and Rich (1994))

One can specify the state-specific effects in two ways - as fixed or random. In the fixed effects approach, the regression intercept is assumed to vary across the states. We then estimate the regression using the least squares dummy variable approach (i.e. using a dummy variable for each state), or using a suitable transformation of the model to facilitate computation. On the other hand, when one estimates using the random effects approach, the state specific effect is modelled as an additional, time-invariant error term for each state. The covariance structure of the composite error term $\alpha_i + \varepsilon_{it}$ allows estimation by the generalised least squares method. This is our preferred specification, as allowing for individual effects is in effect leaving permanent differences in growth rates unexplained. The random effects approach also has an advantage in that it reduces the number of degrees of freedom lost due to the number of dummy variables introduced in the fixed effects approach. However, the random effects approach assumes that the state specific random error is uncorrelated with the other explanatory variables, which may not be the case. Thus to check for the appropriateness of the random effects approach we test for orthogonality of the random effects and the regressors using the HAUSMAN (1978) test We will present results for both fixed and random effects specifications; the results only marginally differ. For all our tests (i.e. tests of significance and the Hausman test), we use the Huber-White estimate of variance, which allows for different error variances across states as well as serial correlation for the states. To account for the endogeneity of the individual infrastructure, we use the method of two stage least squares, to be detailed shortly.

Table 4 tabulates the results. In our first specification (columns 1 and 2) we observe the explanatory power of infrastructure in general, summarised by the indicator INFRA calculated earlier in Section 2, real development expenditure, and the initial level of SDP (in year 1977) with control variables - the share of agriculture in SDP and inflation. Column 1 summarises the fixed effects results, column 2 the random effects. We find that 36 per cent of variation in the growth rates are explained by the first model - this improves marginally for the random effects specification. For both specifications we find the coefficient for infrastructure (the variable used is INFRA, estimated earlier by factor analysis) to be positive and significant. The development expenditure indicator, is not significant in both cases. The coefficient for inflation too is not significantly different from zero in both specifications. The coefficient of the initial level of income is negative, as would be expected, but is not significantly different from zero.

When the state specific effects are specified as fixed, the precision of the estimates decline (i.e. we find that the standard errors increase by about 40%). This is because a great deal of cross section information is absorbed in the state specific dummies. That the coefficients do not significantly differ between random and fixed effects estimates, is confirmed by the Hausman test, where we do not reject the null hypothesis that the state specific effects are orthogonal to the regressors. In other words, one need not reject the random effects model in favour of the fixed effects model.

Columns 3 and 4 present results for a similar specification - only that we replace the general index of infrastructure by some basic infrastructural indicators included individually. The indicators which have a significant influence in explaining inter-state variation in growth rates are the following: percentage of net irrigated area of net cultivated area, per capita industrial power consumption, length of road network per 1000 sq km, infant mortality rate (marginally), primary education, and the ratio of bank deposits to the SDP. The last two indicators can be seen to be proxies for level of education and the depth of the financial sector, respectively. Replacing the variable INFRA by the individual infrastructural indicators increases the explanatory power of the model to almost 40 per cent. All of the indicators are observed to be significant. Our two control variables, the structure of production represented by the share of agriculture, and inflation, do not appear to significantly explain inter-state growth performances. The coefficient for development expenditure, too, is not significantly different from zero.

Irrigation, measured as percentage of gross cropped area irrigated, appears to be a significant explanatory variable in all specifications (including the 2SLS specifications following in columns 5 to 10). The states of Punjab and Haryana are examples of the radical benefits from the Green Revolution implemented in the mid sixties, which involved creating extensive irrigation facilities, alongside radical land reforms and provision of credit institutions. Per capita consumption of industrial power also appears to be consistently significant across all specifications. Other indicators of power consumption, i.e. that of percentage of villages with electricity, and per capita total consumption do not consistently appear as significant explanatory indicators.

The density of the road network, accounting for the effect of transport and communication, is also revealed to have a positive and significant effect in all specifications. Other physical infrastructure variables, for example, number of vehicles per 1000 inhabitants, length of rail network, do not show up as significant variables explaining cross section growth variation. The importance of road networks over that of railroad connections and that of motor vehicles can be accounted for by the different forms of informal road transport characteristic of poor economies connecting the villages, small townships, semi-urban areas, to the urban townships and cities. Despite developed rail connections within and between states, roads still remain the main means of communication between villages and the nearest townships.

Of the three education variables, primary education appears to have played a significant role in explaining differential growth performances across Indian states. In later specifications, (column 7-10) we find that literacy also explains a significant amount of variation. This too can be explained by the nature of economic development in rural and semi-urban areas and townships, where employment is mostly in the tertiary sector, and is largely informal, requiring skills no more than a primary education. Also, rural India, which constitute over 75% of the Indian population , if at all provided with an educational institution, are most likely to have a government sponsored primary school - hence the significant impact of literacy and primary education in the results.

Finally bank offices per 1000 inhabitants, bank deposits and bank credit as a share of GDP, used as indicators of financial development, are also revealed to significantly explain variation in growth rates across the states.

5.1 Accounting for potential endogeneity bias

Much of the insignificance of many of the explanatory variables in our estimations may be attributed to the endogenous nature of the various infrastructure variables. Reverse causality between infrastructure and economic growth (especially GDP per capita levels) may arise due to a number of reasons. Most infrastructural projects involve a substantial fixed cost which cannot be undertaken unless income is higher than a given threshold. It is also likely that new infrastructure is systematically located in areas where firms have more chances of being successful for reasons other than infrastructure availability. Proximity to markets, coastal areas, primary resources and labour can be factors that can attract productive investment.

To avoid biased estimates because of potential endogeneity we run two-stage least squares regressions. Endogenous (and exogenous) variables have been earlier determined by Granger causality tests[6]. We estimate the instruments by running regressions of each endogenous variable on the variables identified as exogenous, and the resulting predicted values serve as their instruments[7]. The first stage regressions of infrastructure equations are presented in Table 5, using random effects specifications. The results of the Column 5 and 6 in Table 4 now presents results of the fixed effects and random effects regression with previous specifications. We do not observe a significant increase in explanatory power, neither a significant change in the values of the estimates.

In the following specifications, we drop control variables of the share of agriculture in SDP and inflation, and real development expenditure, given their insignificance in the previous two specifications. Columns 7 to 10 present the results. In columns 7 and 8 we use the observed values of the variables (while 2SLS IV estimates of these variables are used in Cols. 9 and 10), for both random and fixed effects - all of the variables used in the previous specification are significant in this specification too. A number of other infrastructural variables are also used in this specification, but are not

[6] Results are not presented in paper due to space constraints.

[7] We do not perform a Sargan test for over-identification as the number of exogenous variables are the same as that of the instrumenting equations.

reported in the results due to their insignificance. We yet again observe that literacy appears to be an important variable in explaining cross section variations in growth. We repeat this test with predicted values of the variables from the infrastructural equations - the main results remain unchanged.

Finally, being unable to find suitable instruments for state development expenditure, we use the residuals extracted from the lead-lag regressions, used for the conditioning distribution dynamics as the instrument. The results reveal yet again that development expenditure does not explain any variation in growth rates across states.

To summarise: Unconditional convergence is not obtained in the panel regression approach, while conditional convergence is not obtained. Using a two-stage least squares approach, we find that net area irrigated, primary education, road density, literacy, roads, power consumption in industrial sectors, and bank deposits to be infrastructural components which significantly explain inter-state variation in growth. Of these, the effect of education, the extent of irrigation and roads have been found to be robust across various specifications. The infrastructure index is also found to be significant and positive in effect.

6 Conclusion

This paper has examined the convergence of growth and incomes with reference to the Indian states using an empirical model of dynamically evolving distributions. The dominant stylised facts are of twin-peaked dynamics over the period 1965-1997 - with some cohesive tendencies in the 1960s, only to dissipate over the following three decades. These findings contrast with those emphasised in works of Aiyar (2000), Bajpai and Sachs (1996), Nagaraj and Venganzones (1997), where divergent tendencies are highlighted, but increasing polarisation and club convergence remains undocumented.

I show that economic growth across the Indian states has polarised into two income convergence clubs: one at 50 per cent of the national average, the other at 125 per cent of the national average. The period 1965-1970 exhibits some tendencies of convergence, which are eventually found to dissipate in the seventies, eighties and the nineties.

A conditioning methodology using the same empirical tools further reveals that such income dynamics are explained by the disparate distribution of various infrastructures. Infrastructure is found to explain the formation of only the lower convergence club and not the upper - this would not have been revealed with standard tools of convergence analyses. Of the different indicators, the most convincing results are obtained with education, literacy rates, and percentage of irrigated land, which individually explain the formation of the lower income club, and to a certain extent that of the higher income club. The results obtained are distinctly different from those obtained in earlier studies in that it isolates "conditional convergence" at specific parts of the income distribution - results which would go uncovered in standard methods of investigating for conditional convergence using regression analyses.

Standard parametric conditioning results complement the distribution dynamics results. Of the infrastructure indicators we find the extent of irrigation, education, and in particular literacy, roads, power consumption in industrial sectors, and bank deposits to be infrastructural components which significantly explain inter-state variation in growth. Of these, the effect of education, the extent of irrigation and roads have been found to be robust across various specifications. Conditional convergence is very rarely observed, if at all.

The empirical results suggest that the relationship between infrastructural indicators and economic growth is a significant one - and that this is especially so for the lower income states. It would therefore be useful to have a well-defined model defining the channels through which infrastructure promotes growth.

References

Aiyar, S. (2000). Growth theory and convergence across Indian states: A panel study, in India at the crossroads: Sustaining growth and reducing poverty, edt by T. Callen, P. Reynolds and C. Towe.

Bajpai, N. and J. Sachs (1996). Trends in inter-state inequalities of income in India", Discussion Paper No. 528, Harvard Institute for International Development, Cambridge, Mass., May.

Bell, C. and R. Rich (1994). Rural poverty and agricultural performance in post-independence India. *Oxford Bulletin of Economics and Statistics, 56, 111-33..*

Ben-David, B. (1994). Converging clubs or diverging economies. *Working Paper 922, CEPR , London, February.*

Benhabib, J. and M. Spiegel (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics, 34*, 143–173.

Benhabib, J. and M. M. Spiegel (2000). The role of financial development in growth and investment. *Journal of Economic Growth, December, 5(4)*, 341–360.

Bernaud, A. and S. Durlauf (1994). Interpreting tests of the convergence hypothesis. *NBER Working Paper No. 159*.

Bianchi, M. (1995). Testing for convergence: A bootstrap test for multimodality. *Journal of Applied Econometrics, (21)..*

Bulli, S. (2001). Distribution dynamics and cross-country growth - a new approach. *Scottish Journal of Political Economy 48, 2*, 226–243.

Burgess, R. and R. Pande (2003). Do rural banks matter? Evidence from the Indian social banking experiment. LSE mimeo.

Carlino, G. and L. Mills (1993). Are U.S. regional incomes converging? a time series analysis. *Journal of Monetary Economics*, 335 –346.

Cashin, P. and R. Sahay (1996). Internal migration, centre-state grants, and economic growth in the states of india. *IMF Staff Papers, 43(1)..*

Chung, K. (1960). Markov chains with stationary transition probabilities. Springer-Verlag, New York.

De Long, J. B. (1988). Productivity growth, convergence, and welfare. *American Economic Review: A Comment, 78(5):1138-55, December.*

Desdoigts, A. (1994). Changes in the world income distribution: A non-parametric approach to challenge the neo-classical convergence argument. *PhD dissertation, European University Institute, Florence.*

Durlauf, S. and P. Johnson (1994). Multiple regimes and cross country growth behaviour. *Working Paper, University of Wisconsin, May..*

Esteban, J. and D. Ray (1994). On the measurement of polarisation. *Econometrica, 62(4)*, 819–851.

Fiaschi, D. Lavezzi, A. (2003). Distribution dynamics and nonlinear growth. *Journal of Economic Growth, Springer vol. 8(4)*, 379–401.

Galor, O. and D. Tsiddon (1994). Human capital distribution, technological progress, and economic growth, CEPR working paper no. 971.

Galor, O. and J. Zeira (1993). Income distribution and macroeconomics. *Review of Economic Studies*, 35–52.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica 46(6)*, 1251–71.

Jones, C. I. (1997). On the evolution of the world income distribution. *Journal of Economic Perspectives, 11(3)*, 19–36.

Kalaitzidakis, P., T. M. and A. Savvides (2001). Measures of human capital and non-linearities in economic growth. *Journal of Economic Growth 6*, 229–254.

King, R. G. and R. Levine (1992). Financial indicators and growth in a cross section of countries. *World Bank Policy Research Working Paper No. 819*.

Lee, R. J. B. J.-W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics, Vol. 32, no. 3*, 363–394.

Loewy, M. and D. Papell (1994). Are US regional incomes converging: Some further evidence. *mimeo, University of Houston.*

Nagaraj, R, A. V. and M. Venganzones (1997). Long run growth trends and convergence across indian states. *Technical paper no.131, OECD Development Centre, Paris..*

Özler, B, G. D. and M. Ravallion (1996). A database on poverty and growth in india. *Poverty and Human Resources Division, Policy Research Department, The World Bank.*

Quah, D. (1996). Convergence empirics across economies with (some) capital mobility. *Journal of Economic Growth, March 1996, vol. 1, no. 1*, 95–124.

Quah, D. (1997). Empirics for growth and distribution: Stratification, polarisation and convergence clubs. *Journal of Economic Growth, 2(1)*, 27–59.

Solow, R. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics, 70(1)*, 65–94.

Stokey, N. L. (1991). Human capital, product quality and growth. *Quarterly Journal of Economics, Vol. 106*, 587–616.

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

7 Data Appendix

States used in the study:

Andhra Pradesh
Assam
Bihar
Delhi
Gujarat
Haryana
Jammu and Kashmir
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Orissa
Punjab
Rajasthan
Tamil Nadu
Uttar Pradesh
West Bengal

Other states were excluded from the study due to the incomplete data available over the given period. These states together constitute for over 80% of the national population.

Price data that has been used to deflate the nominal GDPs has also been obtained from the above mentioned data set, and is the adjusted CPIAL index.

State development expenditure constitutes of expenditure on both economic and social services. The economic services include agriculture and allied activities, rural development, special area programmes, irrigation and flood control, energy, industry and minerals, transport and communications, science technology and environment; the social services include education, medical and public health, family welfare, water supply and sanitation, housing, urban development, labour and labour welfare, social security and welfare, nutrition, and relief on account of natural calamities.

8 Appendix

Quah (1997) exploits a duality property from Markov process theory to provide a tractable model of distribution dynamics. To model the distribution dynamics, one observes a scalar stochastic process, and then derives the implied unobservable sequence of distributions associated with this process. This hypothesised distribution sequence is then defined to be the dual to the observed scalar stochastic process. The property is reversed (the mathematics involved, however, remaining unaffected) to track the distribution dynamics as follows: while the sequence of distributions is observed, its dual, the scalar stochastic process, is implied, though unobserved. The dynamics of the scalar process is described in a transition probability matrix, while the dual to this, the stochastic kernel, describes the "law of motion" of the sequence of distributions. These will serve as models which describe the distribution dynamics across the Indian states.

The following clarifies the concepts discussed above. Let F_t be the measure corresponding to the cross-country income distribution at time t. The stochastic kernel which measures the evolution from F_t to F_{t+1} is a mapping M_t from the Cartesian product of income values and Borel measurable sets to $[0, 1]$, such that

$$\nabla Borel - measurable A, F_{t+1}(A) = \int M(A, y) dF_t(y) \tag{2}$$

It is M_t which encodes all the information about the evolution, or the law of motion of the sequence of distributions over time periods t and $t + 1$. It contains information of the intra-distributional dynamics, hence revealing specific external shapes of the distribution, unrevealed in standard empirical procedures. M_t is assumed to be time-invariant, (and in this case, leaving out an error term, inclusion of which would render the model as analogous to a first order vector auto-regression in distributions rather than scalars or finite dimensional vectors), one can re-write the above expression as

$$F_{t+1} = MF_t \tag{3}$$

For simplicity in calculations, iterating the above equation and leaving out the error term, one can write:

$$F_{t+s} = M^s F_t \tag{4}$$

As $s \to \infty$ it is possible to characterise the long run distribution - this is called the ergodic distribution and it predicts the long term behaviour of the underlying distribution. If F_{t+s} degenerates to a point mass one can conclude that there is a tendency to global convergence. If F_{t+s} tends towards a bi-modal distribution (the case with the Indian states) one can conclude that there tendency to polarization, with the rich and the poor being pulled apart. Different variants of equation (1) allow the researcher to derive the various spectral characteristics of M_t, such as intra-distributional mobility and the speed of convergence.

Table1a: Inter-State (per capita) income dynamics, 1965-70

First Order transition matrix, Time stationary

(Number)	Upper end point				
	0.640	0.761	0.852	1.019	1.393
5	0.40	0.00	0.40	0.00	0.20
5	0.00	0.40	0.20	0.20	0.20
2	0.00	0.00	0.50	0.00	0.50
4	0.00	0.00	0.25	0.25	0.50
1	0.00	0.00	0.00	1.00	0.00
Ergodic	0.00	0.00	0.22	0.44	0.33

Table1b: Inter-State relative (per capita) income dynamics, 1971-80

First Order transition matrix, Time stationary

(Number)	Upper end point				
	0.680	0.730	0.795	1.010	1.489
5	0.40	0.60	0.00	0.00	0.00
1	0.00	1.00	0.00	0.00	0.00
3	0.00	0.67	0.33	0.00	0.00
4	0.00	0.00	0.75	0.25	0.00
4	0.00	0.00	0.00	0.50	0.50
Ergodic	0.00	1.00	0.00	0.00	0.00

Table1c and d: Inter-State relative (per capita) income dynamics, 1981-89 and 89-97

First Order transition matrix, Time stationary

Number	Upper end point				
	0.533	0.628	0.795	1.010	1.489
6	0.17	0.50	0.33	0.00	0.00
4	0.00	0.00	0.25	0.75	0.00
3	0.00	0.67	0.33	0.67	0.00
2	0.00	0.00	0.00	0.00	1.00
2	0.00	0.00	0.00	0.00	1.00
Ergodic	0.00	0.00	0.00	0.00	1.00

(Number)	Upper end point				
	0.141	0.207	0.241	1.412	1.489
5	0.40	0.60	0.00	0.00	0.00
1	0.00	1.00	0.00	0.00	0.00
3	0.00	0.67	0.33	0.00	0.00
4	0.00	0.00	0.75	0.25	0.00
4	0.00	0.00	0.00	0.50	0.50
Ergodic	0.00	1.00	0.00	0.00	0.00

Table 2

Results of Factor Analysis

Components	Eigenvalue	Cumulative R^2
f1	12.41	0.83
f2	1.22	0.91
f3	1.00	0.97

Factor Loadings

	f1	f2	f3
total power consumption	0.97	-0.16	0.10
power consumption in industrial sector	0.95	-0.12	0.04
percentage of villages electrified	0.99	0.04	-0.08
percentage of net area operated with irrigation	0.95	-0.20	0.18
length of road network per 1000 sq kms.	0.97	-0.12	0.10
number of motor vehicles per 1000 inhabitants	0.89	0.07	-0.37
length of rail network per 1000 sq.kms	0.61	-0.47	0.60
literacy rate of adult population	0.98	-0.04	-0.15
primary school enrolment rate	0.97	0.04	-0.08
secondary school enrolment rate	0.98	-0.13	-0.02
infant mortality rate	-0.96	0.05	0.22
bank offices per 1000 people	0.91	0.24	-0.30
bank deposits as a percentage of SDP	0.75	0.57	0.28
bank credit as a percentage of SDP	0.58	0.68	0.40

Table 3. Conditioning regressions (two sided projections) of growth rates on infrastructure

infrastructure		Co-efficients in two-sided projections		
Lead	4			-0.005 (0.003)
	3		0.01 (0.01)	0.012 (0.004)
	2	0.03 (0.04)	-0.02 (0.03)	-0.03 (0.021)
	1	0.065 (0.01)	0.068(0.012)	0.07 (0.02)
	0	0.072 (0.016)	0.074 (0.018)	0.079 (0.012)
Lag	1	-0.071 (0.014)	-0.067 (0.016)	-0.069 (0.012)
	2	-0.042 (0.010)	-0.04 (0.011)	-0.033 (0.011)
	3		-0.014 (0.012)	-0.008 (0.005)
	4			-0.005 (0.004)
Constant				
Sum of co-efficients		-0.054	0.011	0.011
R^2		0. 17	0. 15	0. 12

Note: Figures in parentheses are white heteroskedasticity consistent standard errors.

Table 4. Panel Regressions

dependent variable $\ln y_t - \ln y_{t-1}$	1 FE	2 RE	3 FE	4 RE	5 FE – IV	6 RE-IV	7 FE	8 RE	9 FE-IV	10 RE-IV
initial income level	-0.015 (.479)	-0.015 (0.388)	-0.013 (0.65)	-0.013 (0.34)	-0.013 (0.72)	-0.013 (0.78)	-0.016 (0.71)	-0.016 (0.88)	-0.017 (0.715)	-0.017 (0.9)
share of agriculture in SDP	-0.17 (0.69)	-0.17 (0.62)	-0.17 (0.75)	-0.17 (0.55)	-0.17 (0.69)	-0.15 (0.72)				
Inflation	-.218 (1.27)	-0.217 (1.25)	-0.2 (0.21)	-0.2 (0.22)	-0.21 (0.14)	-0.20 (0.22)				
Index of infrastructure (f1)	0.001* (4.78)	0.001* (4.74)								
state development expenditure	-.02 (0.172)	-0.02 (0.146)	-0.05 (0.33)	-0.05 (0.26)	-0.05 (1.78)	-0.05 (1.92)				
%net irrigated area of net cultivated area			0.169* (5.87)	0.168* (5.99)	0.169* (6.01)	0.169* (6.99)	0.178* (6.39)	0.178* (6.39)	0.178* (6.01)	0.179* (6.39)
per capita industrial power consumption			0.021* (6.99)	0.022* (7.21)	0.02* (5.12)	0.02* (5.6)	0.062* (9.48)	0.063* (9.6)	0.063* (8.29)	0.063* (8.28)
length of road network per 1000 sq.			0.033* (4.99)	0.003* (5.59)	0.003* (6.01)	0.003* (7.89)	0.004* (8.19)	0.004* (8.67)	0.004* (6.29)	0.004* (6.99)
Literacy of adult population							0.485* (6.86)	0.487* (6.99)	0.485* (6.28)	0.485* (6.43)
Primary school enrolment rate			0.073* (4.23)	0.072* (5.17)	0.062* (5.01)	0.063* (5.13)	0.086* (10.01)	0.087* (10.12)	0.085* (10.06))	0.084* (10.11)
Infant mortality rate			-0.007** (2.25)	-0.007** (2.02)	-0.006** (2.12)	-0.006** (2.45)				
bank deposits as a % of SDP			0.012* (5.2)	0.012* (4.3)	0.012* (4.4)	0.012* (4.9)	0.106* (4.71)	0.106* (4.89)	0.106* (4.72)	0.106* (4.8)
R^2	0.36	0.36	0.39	0.39	0.39	0.39	0.64	0.64	0.64	0.64
Hausman specification test		8.4 (0.39)		8.2 (0.43)		8.7 (0.31)		9.3 (0.28)		9.2 (0.30)

Notes: 1. Absolute t ratios in parentheses
2. * denotes that coefficient is significantly different from zero at 5% level, ** at 10% level.
3. The Hausman test is a test of random vs fixed effects
4. IV equations are based on predicted values of variables summarised in Table 5.

Table 5. Infrastructure Equations

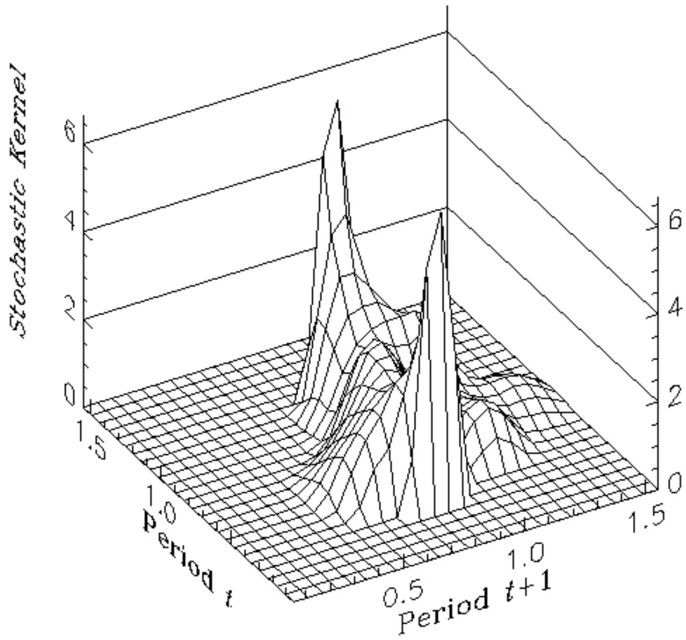
	1 power conspn	2 length of roads	3 primary enrolment	4 infant mortality rate	5 bank deposits
Share of agriculture in SDP			-0.55 (0.72)	-0.55 (0.92)	
Share of industry-transport in SDP	0.17 (2.34)	0.58 (4.28)	-1.03 (8.36)		0.388 (3.03)
Percentage of villages electrified	0.09 (5.83)	0.48 (3.26)	0.39 (11.74)	-0.39 (11.6)	0.09 (5.9)
Length of rail network	36.12 (15.65)	6.49 (11.2)		-0.23 (2.84)	3.27 (8.71)
Percentage of population literate	2.12 (8.93)		-0.64 (7.94)	-1.03 (8.94)	0.42 (15.0)
Percentage of popn. with sec'dary education	4.85 (17.67)				
No. of banks in area per 1000 inhabitants		0.39 (6.37)	1.56 (7.16)		0.14 (8.99)
R^2 adjusted	0.97	0.97	0.98	0.98	0.96
Number of observations	255	255	255	255	255
Hausman specification test (p values)	9.22 (0.29)	12.7 (0.32)	8.72 (0.40)	6.7 (0.42)	8.9 (0.39)

Notes: 1. Absolute t ratios in parentheses

2. * denotes that co-efficient is significantly different from zero at 5%, ** at 10%

3. Hausman test is a test of random vs. fixed effects

Fig.1a and b: Relative Income Dynamics across Indian States, 1 year horizon, 1965-70 and 1971-80



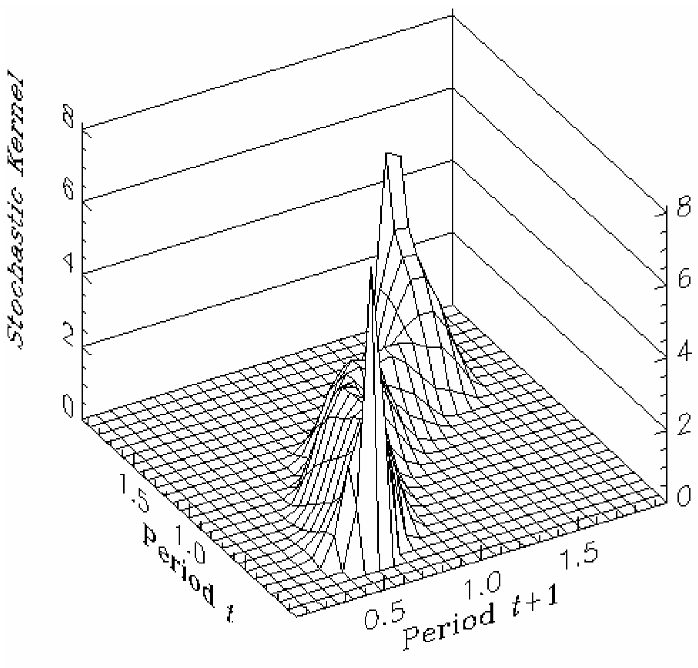
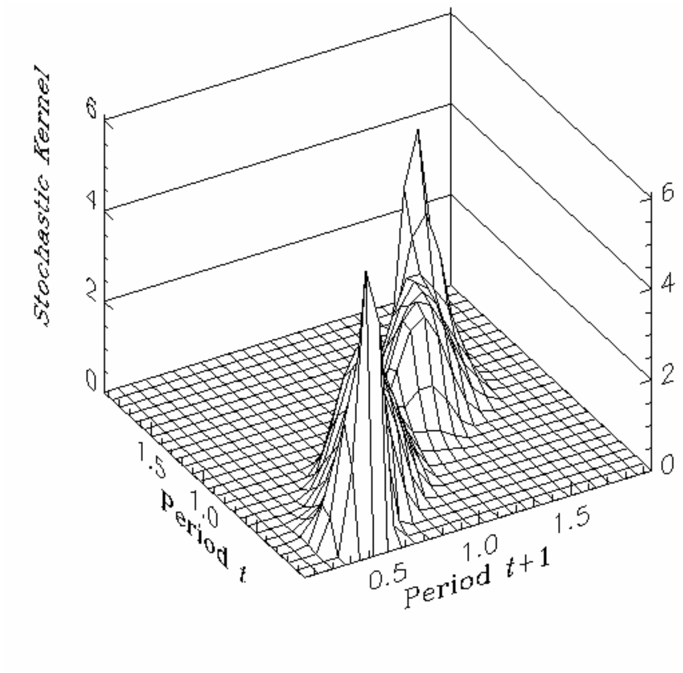


Fig. 1c and d: Relative Income Dynamics across Indian States, 1 year horizon

1981-89 and 1990-97



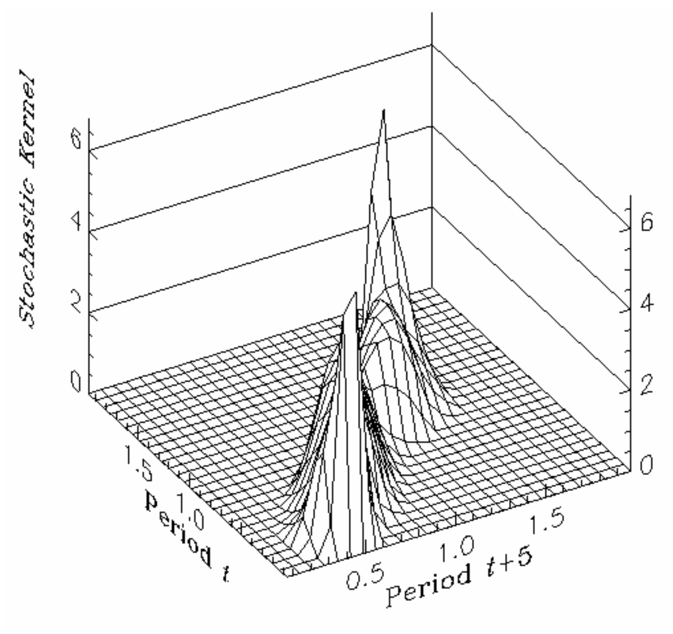


Fig2. Infrastructure dynamics across Indian states
Contour plot, 1978-1993

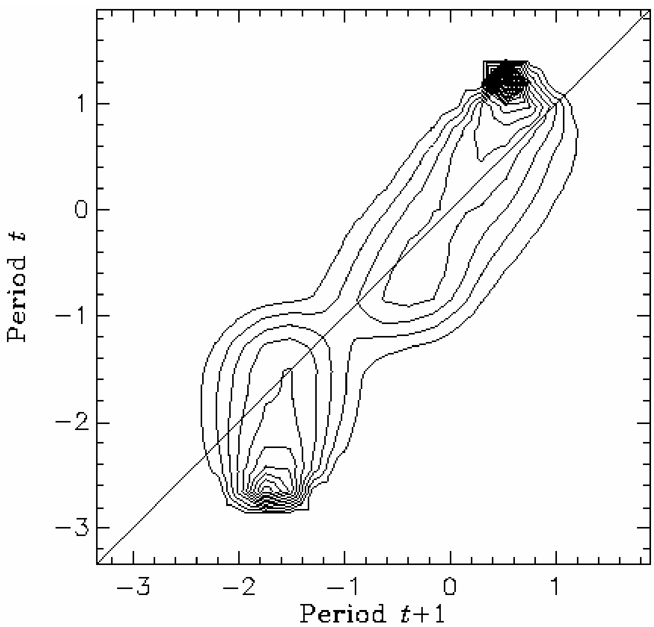
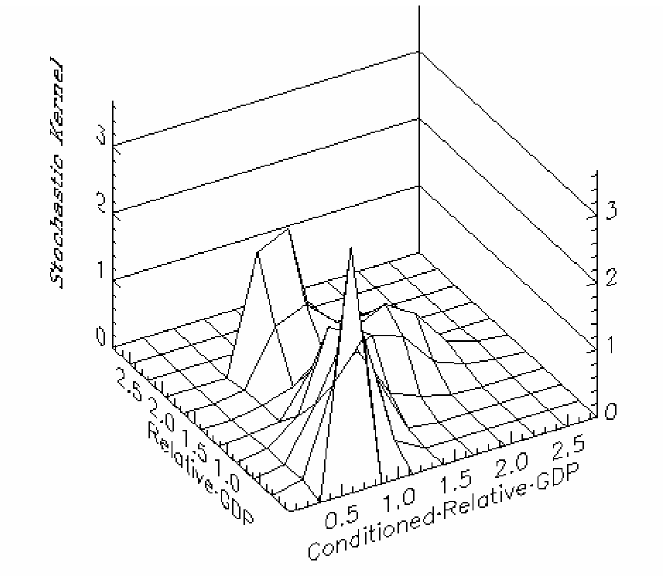


Fig.3a and b Relative per capita incomes across Indian states

Infrastructure conditioning, with contour plot



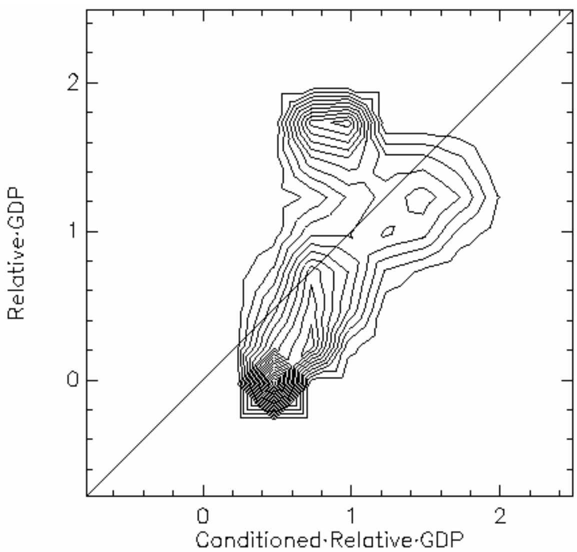


Fig 4a and b. Relative per capita incomes across Indian states
Education conditioning, with contour plot

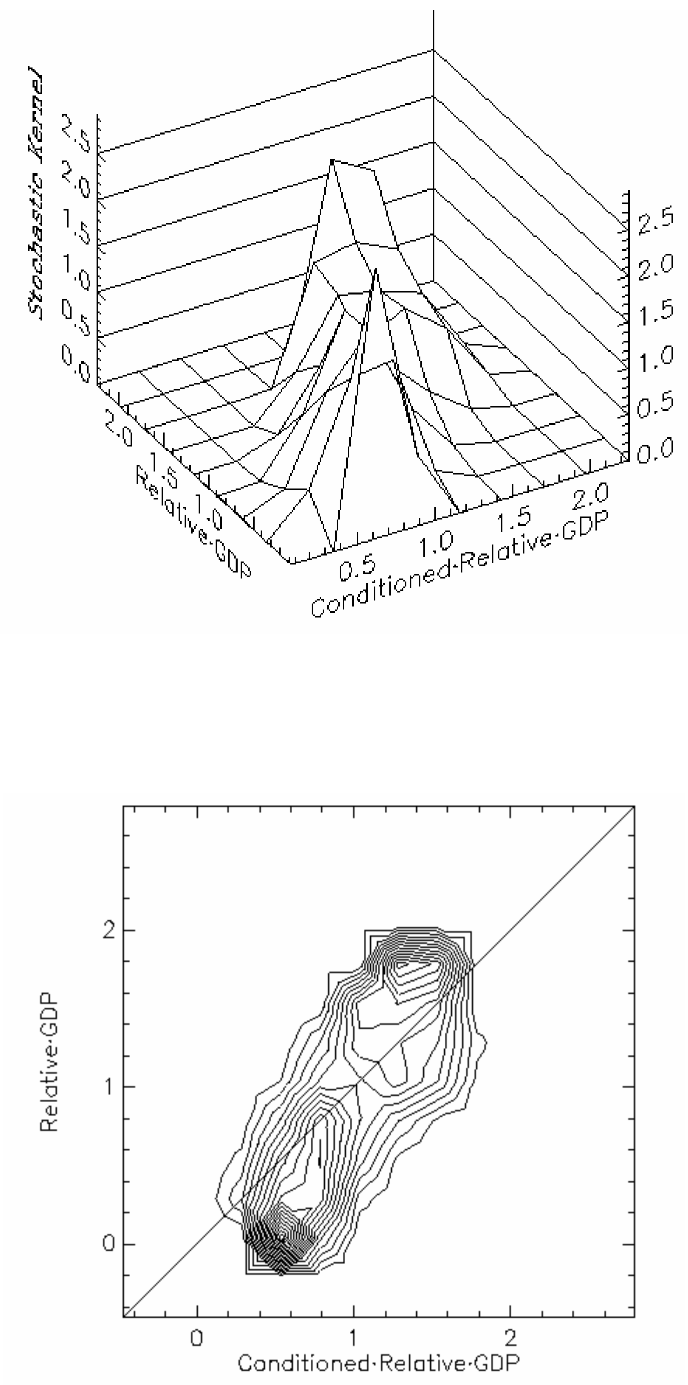


Fig.5. Relative per capita incomes across Indian states

Percentage of Population Literate

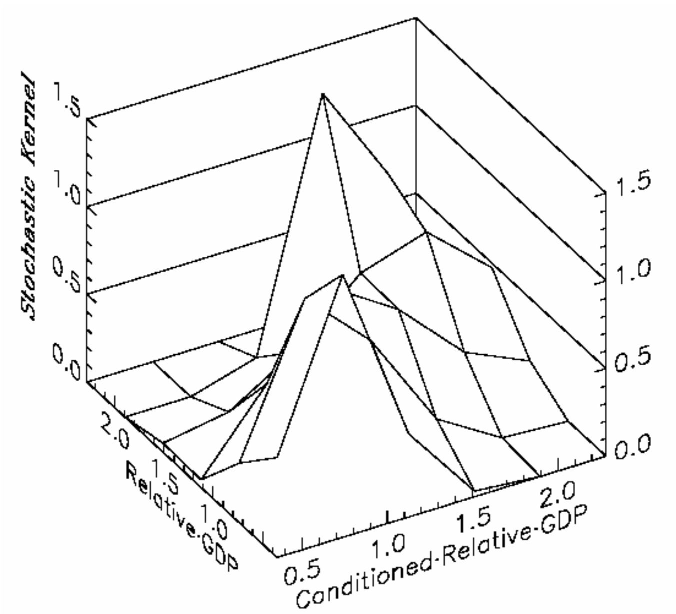


Fig.6. Relative per capita incomes across Indian states

Percentage of Irrigated land conditioning

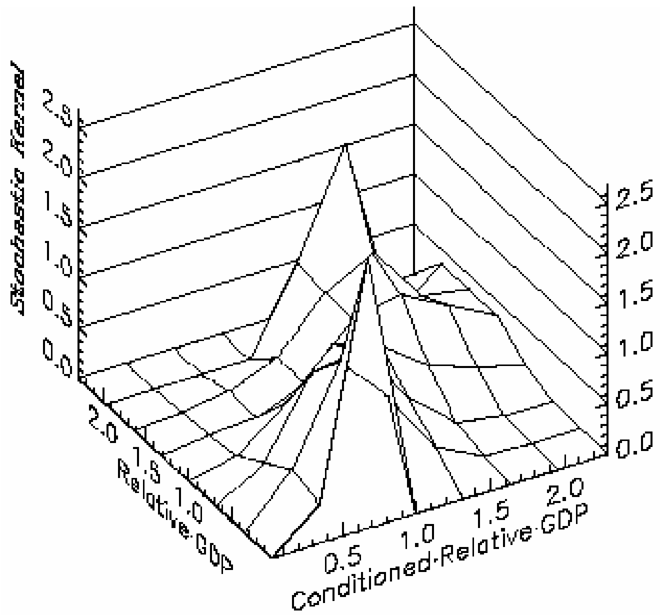


Fig.7. Relative per capita incomes across Indian states

State development expenditure conditioning