Ghate, Chetan; Wright, Stephen

Working Paper

The V-factor: distribution, timing and correlates of the great Indian growth turnaround

DIW Discussion Papers, No. 783

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Ghate, Chetan; Wright, Stephen (2008) : The V-factor: distribution, timing and correlates of the great Indian growth turnaround, DIW Discussion Papers, No. 783, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
https://hdl.handle.net/10419/27307

Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

Discussion Papers

783

Chetan Ghate • Stephen Wright

The "V-Factor": Distribution, Timing and Correlates of the Great Indian Growth Turnaround

Berlin, April 2008

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the institute.

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Available for free downloading from the DIW Berlin website.

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN.
Papers can be downloaded free of charge from the following websites:

http://www.diw.de/english/products/publications/discussion_papers/27539.html
http://ideas.repec.org/s/diw/diwwpp.html
http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1079991

The "V-Factor": Distribution, Timing and Correlates of the Great Indian Growth Turnaround*

Chetan Ghate† and Stephen Wright‡

Abstract

The ratio of Indian to US per capita output over the past 45 years has displayed a distinctive "V"-shaped pattern. We show that a strikingly similar V-shaped pattern is visible not just in aggregate output figures, but also as the primary determinant of long-term movements in the cross-sectional distribution within the All-India total, at both sectoral and state output levels. We also carry out preliminary investigations of correlates of the "V-Factor", using a new panel data set for Indian states from 1960 to 2005 that extends and encompasses all previous datasets relevant to macroeconomic analysis of the Indian states.

JEL classifications: O10, O40, O53, O47

Keywords: Principal Components, Divergence, Indian States.

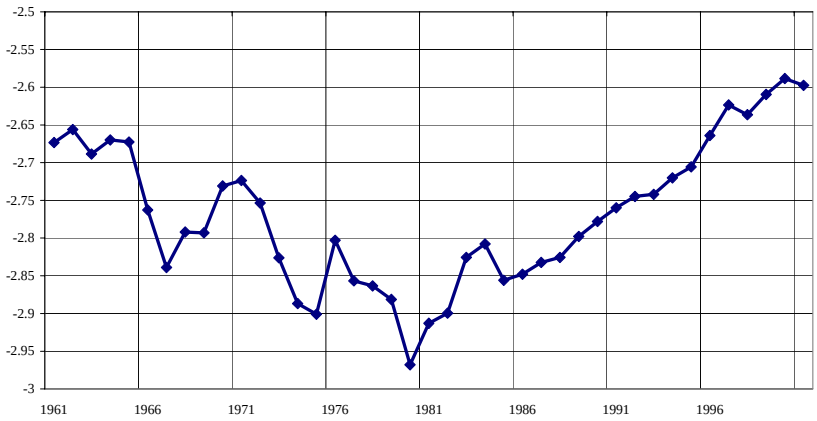
*We thank Amit Sadhukhan for invaluable research assistance. We thank Dr. Savita Sharma and Pronab Sen of the CSO for helpful advice on the data. We also thank seminar participants at ICRIER for useful comments. Stephen Wright is grateful to the Indian Statistical Institute, Delhi, for hospitality during research visits in 2007 and 2008.

†Indian Statistical Institute, Delhi, and German Institute of Economic Research (DIW), Berlin

‡Department of Economics, Birkbeck College, University of London, Malet Street, London W1E 7HX, UK. s.wright@bbk.ac.uk

Figure 1

log GDP per Capita: India vs USA



1 Introduction

The central themes of this paper are best illustrated by a single chart. Figure 1 shows that the ratio of Indian to US per capita output over the past 45 years has displayed a distinctive "V"-shaped pattern. Until the 1980s India's output growth was systematically lower than that of the US; subsequently it has been systematically higher. Taken as a whole India has now been clearly converging for at least two decades.

This remarkable turnaround in growth has not been uniformly distributed, whether across sectors or across states. However, we show in this paper that a strikingly similar V-shaped pattern is visible not just in aggregate output figures, but also as the primary determinant of long-term movements in the cross-sectional distribution within the All-India total, at both sectoral and state output levels. Following Bai (2004) and Bai and Ng (2002; 2004) we identify common factors determining long-term growth performance by Principal Components analysis of log output levels. This displays two clearly dominant factors: a common long-term growth factor and a common "V-factor", that appear jointly to capture the permanent components of output per capita, disaggregated both by state and by major industry group over the same period. The common "V-Factor" broadly resembles the series shown in Figure 1. Using these two factors we can identify with some degree of confidence both "V-States" and "V-Sectors": i.e, those

with a positive loading on the V-Factor.

Our analysis has three major strands:

1. We carry out a principal components based sectoral growth accounting exercise that focuses on the contributions of the first two principal components to the long-term growth performance of the Indian States both at the aggregate and sectoral level. One key conclusion we draw is that "V-States" have not been systematically either richer or poorer than Non-V states. Unsurprisingly, the V-States have been those with a preponderance of V-sectors. In line with other past research, we find that the most marked "V"s are visible in the service sector; but manufacturing has also made important V-factor contributions in some states. In general, the V-factor makes smaller contributions to long-term growth of agriculture.

2. Our analysis also casts some light on the debate on when the turn-around in growth took place. The aggregate figure shown in Figure 1 suggests (roughly in line with the findings of Rodrik and Subramanian (2005) and others, that the low-point of the V was right at the start of the 1980s - at a time which is hard to rationalize in terms of policy changes. In conrast our more disaggregated analysis typically produce estimates of the V-Factor that have a low-point distinctly later in the 1980s, and thus are more readily explicable in terms of policy changes. However, we would be cautious in drawing too precise an inference on this issue. The data provide much stronger evidence for the existence of a long-term V-factor than they do for the precise location of its apex, which can appear to shift by several years simply on the basis of which particular dataset is used.

3. A final, more preliminary and speculative component of our analysis relates to the nature of the V-Factor itself. Its ubiquity suggests a strong common element, which it seems reasonable to ascribe to policy changes. But the open question is why policy changes that were common across states appear to have had such uneven, and sustained uneven effects. We certainly do not claim to have found a causal explanation, but can at least point to two striking correlations with the V-Factor. The first is that the public sector, whether defined in terms of output, or in terms of development expenditures, shows

strong evidence of being negatively correlated with the V-Factor, i.e., is "Anti-V" This is consistent with the findings of some earlier research.[1] The second is the role of supply constraints which may have limited the impact of the V-Factor in some states. Both features of the data suggest that differential loadings on the V-Factor do not reflect movements in long-term conditional steady-state levels of output; but rather in differential levels of frictions that impede movements towards that steady-state.

The structure of the paper is as follows. In Section 2 we relate our analyis to recent literature; Section 3 summarises the key features of the dataset. In Sections 4 and 5 we derive estimates of the V-factor from the method of principal components applied to statewise output both at total and sectoral levels. In Section 6 we contrast the timing of the turnaround in our estimated V-factors with the timing suggested in past research. In Section 7 we provide some preliminary evidence of correlates of the V-Factor; and in Section 8 we attempt to reconcile the V-factor representation with a benchmark model of convergence. Section 9 concludes the paper, while appendices provide additional background detail.

2 Related Literature

This paper can be related to two broad strands of past research. The first examines the sources and timing of the shift in Indian output growth since the 1980s; a second examines the longer-term issue of convergence between the Indian states.

There has been a recent surge of research on the timing and proximate causes of structural breaks in India's growth rate. This literature addresses a range of questions, such as: When was the shift in growth? Was it policy driven? If so, what were the crucial policy changes that drove growth? Was the shift uniform across states? Our results shed light on some of these questions.

Virmani (2006) finds that the upward break in growth in the manufacturing sector is responsible for the structural break in growth. In particular,

[1] Clark and Wolcott (2003) also show that several measurable dimensions of state policy such as the number of phones per 100 workers, the number of kilometers of roads per 100 workers, public education expenditure per 100 workers, or public capital expenditure per 100 workers show little sign of connection with economic growth.

he finds that the growth rate of manufacturing accelerates after 1980-81. This contributes to the acceleration in growth of GDP growth from 1981-82 (p. 92). Virmani also finds no additional breakpoints in the nineties, once the breakpoint in 1980-81 is accounted for.[2] There are two aspects of his results that are worthy of note. First, Virmani finds that 45% of the variation in India's growth rate is explained by fluctuations in rainfall. He uses this fnding to net out the effect of rainfall variation on GDP growth.[3] Second, he uses a standard Chow test to determine structural breaks. However, as noted by Hansen (2001), this approach is problematical since the break-date must be known in advance to the researcher. In the case that the researcher picks an arbitrary candidate break-date, the true break-date can be missed. If the researcher picks a break-date based on some known feature of the data, the Chow test is misleading since the candidate break-date is endogenous.

Balakrishnan and Parameswaran (2007) utilize the approach developed by Bai and Perron (1998) and Bai (2003) which allows for the simultaneous and endogenous estimation of break-dates. In contrast to Virmani (2006), Balakrishnan and Parameswaran (2007) find that the break in growth the rate of GDP occurs in 1978-79 –with the 1978-79 take off in growth occurring prior to the positive break in manufacturing (1982-83). This suggests that the evidence for manufacturing having served as a primary engine of growth through appropriate market reforms is weak.[4]

Rodrik and Subramanian (2005) argue – in similar vein to Virmani (2006) – that the improvement in India's economic performance was driven by policy changes. In particular, Rodrik and Subramanian argue that the trigger for India's upward break in growth – which they pin down to around 1980 – occurs because of an "attitudinal shift" on the part of the national government in 1980 in favor of businesses. They distinguish between attitudinal changes that are pro-business versus pro-market. Pro-market changes favor entrants and consumers by removing impediments to markets. Pro -

[2] He also finds that the acceleration in growth of GDP from services is a gradual process from 1980-81 to 1985-86.

[3] We shall incorporate similar rainfall adjustments in the next draft of this paper. However preliminary investigations suggest that the longer-term nature of our approach makes our results largely invariant to these adjustments.

[4] Balakrishnan and Parameswaran (2007) account for this finding by arguing that only registered manufacturing breaks in 1982-83, while unregistered manufacturing breaks only in the mid 1980's. Because the share of registered manufacturing in GDP is small (8.7% in 1982-83), they argue that this is unlikely to serve as an engine of growth.

business changes focus on raising the profitability of industrial and commercial establishments and favors incumbents and produces. This shift increased overall productivity. Rodrik and Subramanian use three measures related to aggregate growth performance (real GDP per capita, real GDP per worker, and total factor productivity (TFP)). Each of these variable displays a sharp upward trend beginning in 1979. They also present simple Barro style cross country growth regressions for the periods 1960 – 1980 and 1980 – 99. The TFP regressions show that after controlling for policies, endowments, and initial income, Indian grew 2.1% faster than the average country in the 1980-99 period. They also show that the Indian TFP experience in 1980 - 89 has surpassed that of East Asia even in the first twenty years of the East Asian miracle.

Other authors such as Wallack (2003) also support these findings. In particular, Wallack (2003) finds evidence for a break in the GDP growth rate in the early 1980's. This is close to the result reported by Rodrik and Subramanian (2005). Finally, Hausmann et al. have analyzed transitions to higher growth in a large cross national sample, and date the Indian growth break to 1982.

In Section 6 we show that our more disaggregated approach results in alternative (and distinctly later) estimates of the turnaround in growth, without resorting to any assumptions about exogenous break-points.

Although our paper is only tangentially related to the large literature on convergence in Indian states (see for example Dasgupta et al. (2000), Datt and Ravallion (2002) and Trivedi (2002)), it does provide some insights. Several previous researches have concluded that there is evidence for conditional convergence; but we suspect these results are due to biases in panel estimation techniques. In Section 3.2 we present some simple graphical evidence against both conditional and unconditional convergence; in Section 4 we show how our V-Factor analysis can be reconciled with this result; and in Section 8 we discuss how our results can be consistent with convergence of some states towards the global frontier, but lack of convergence between the Indian states.

3 Some key features of the data.

3.1 The dataset

Our core dataset is a balanced panel of statewise real net domestic product per capita for the 16 major Indian states, for all of which we have constructed continuous series on an annual basis from 1960 to 2003.[5] The 16 major states that we analyse constitute 97% of the Indian population. We also have a breakdown of the statewise total into 13 major sectors, also measured on a real per capita basis. At present we only have a balanced panel of sectoral level data from 1970 onwards. In a future version of this paper we plan to extend these data at least back to the mid-1960s. The NSDP as well as the sectoral data are from the Economic and Political Weekly (EPW) Research Foundation statewise data set. Since the EPW research foundation does not convert all data into a common base year, we have spliced all the NSDP as well as sectoral data so that they are in 1993-1994 prices. We have also corrected for changes in state definitions.

We also have data on a wide range of regional indicators on a statewise basis. Some of these, such as population, literacy, urbanization, are drawn from census data and hence are only available on a decennial basis; but we also have some other true time series data from at least the 1970s onwards. To the best of our knowledge, our dataset as a whole extends and encompasses all previous datasets relevant to macroeconomic analysis of the Indian states (see, Ozler, Datt, and Ravallion (1996), Besley & Burgess (2000)). This is in itself one of the novel contributions of this paper. A full description of the dataset is given in the Appendix. We aim to make the dataset publicly available in the near future.

3.2 Convergence between the Indian states?

To focus the analysis of the paper, we only briefly summarise the nature of the evidence on convergence between the Indian states.. It might be expected that unconditional convergence would be easier to observe at the regional level than at the international level because of similarities in preferences and technology and the basic institutional and political environment. Indeed, Barro & Sala-i-Martin's (1992) original empirical results suggested

[5] We have recently acquired data to 2005; these will be incorporated into empirical work in the next draft of this paper. See Appendix for a fuller description of the dataset.

strong evidence of long-term unconditional convergence for the US states and (somewhat less strongly) for EU countries. However, we find that for Indian states in our sample the evidence is both strongly against unconditional convergence and almost equally strongly against conditional convergence to a fixed point in the cross-sectional distribution of income.

Figure 2

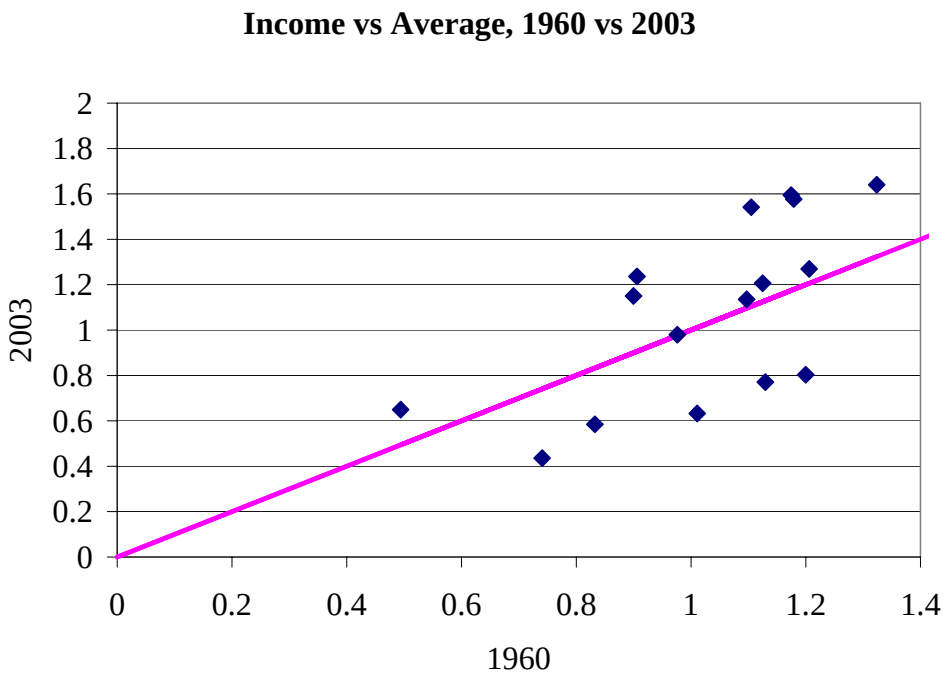


Figure 2 illustrates the evidence against either form of convergence. It plots real state income levels, relative to the average, in 1960 against the same relative values in 2003. The chart also shows a line with unit slope as a basis for comparison. Unconditional convergence would imply that the scatter of points would tend to lie on a line with slope less than 1, since all states would be expected to converge towards the average. Conditional convergence to a fixed point in the distribution would suggest that all points should lie roughly on the line with unit slope. If anything the actual distribution of points appears to lie on a line with slope *greater* than unity, implying that Indian states showed some tendency to *di*vergence over this period.[6]

[6] These results are at odds with some recent research, which has tended to find evidence

3.3 The time-profile of statewise output per capita.

Figure 3. Log Real Output Per Capita in the 16 Major Indian States

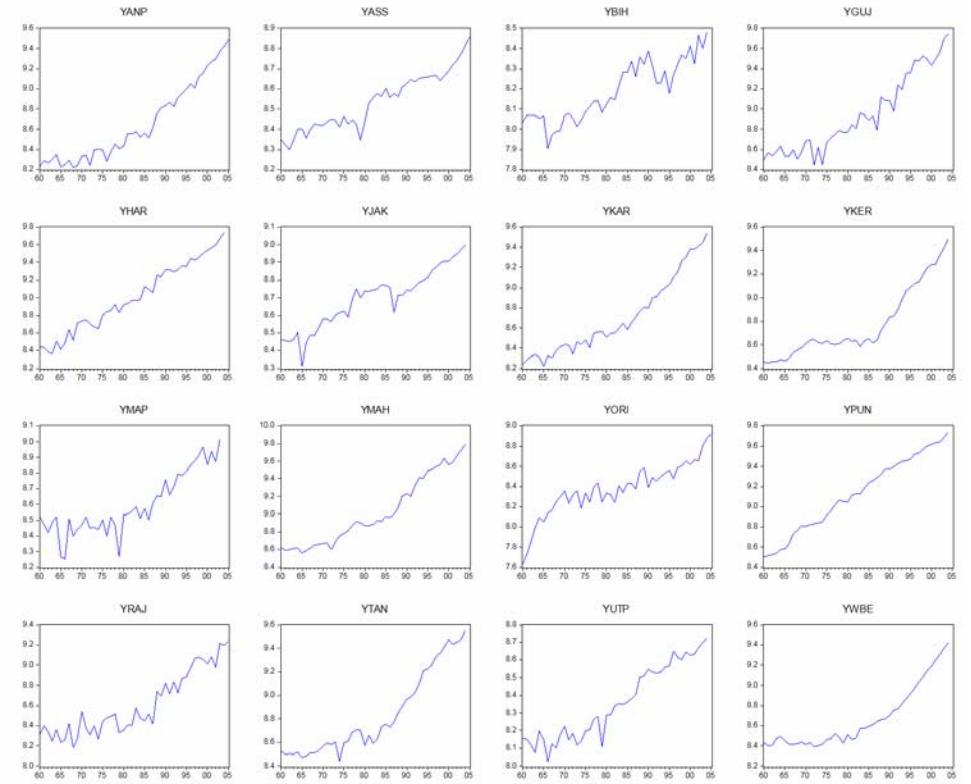


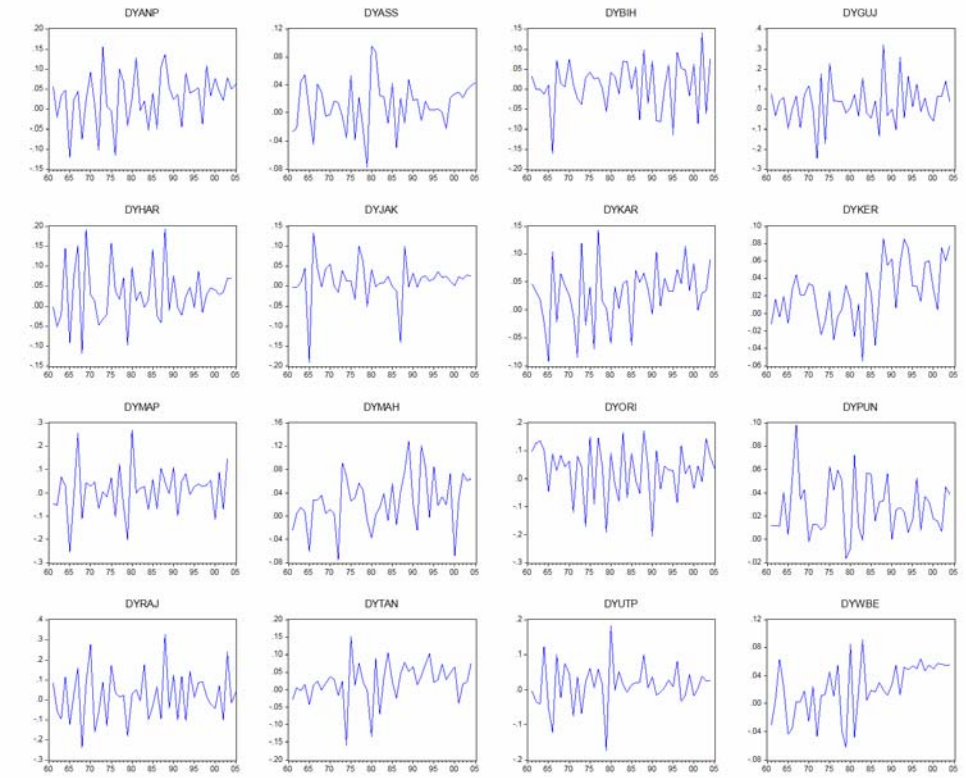
Figure 2B shows total output per capita figures in logs (y) for the sixteen states from 1960 to 2003. Even on the basis of visual inspection there appear to be clear differences in the pattern over time. Some states display what appears to be a clear break in trend near the mid-point of the sample. The states appearing to show such a break are very diverse in nature: Kerala (KER), Madhya Pradesh (MAP), Gujarat (GUJ) and Rajasthan (RAJ) all appear to show a clear break. In contrast, an equally diverse group of states appears to show no obvious break: for example, Bihar (BIH), Haryana (HAR), Orissa (ORI) and Punjab (PUN).

Because log levels are usually assumed to be non-stationary, Figure 4 plots the log changes of real NSDP per capita. This provides a useful reminder that in this form, which should at least be much closer to stationarity,

for conditional convergence (see, for example, Dasgputa et al. (2000), and Trivedi (2002)). However these results, largely based on panel estimation techniques, do not appear to correct sufficiently for known downward biases in coefficients on lagged dependent variables in panels, which in turn tend to overstate the significance of rates of convergence.

short-run volatility is such that longer-term shifts appear much less obvious to the naked eye than those that appear to be present in the levels data. Nonetheless even in the differenced data upward shifts in average growth rates do appear to be visible in a number of states.

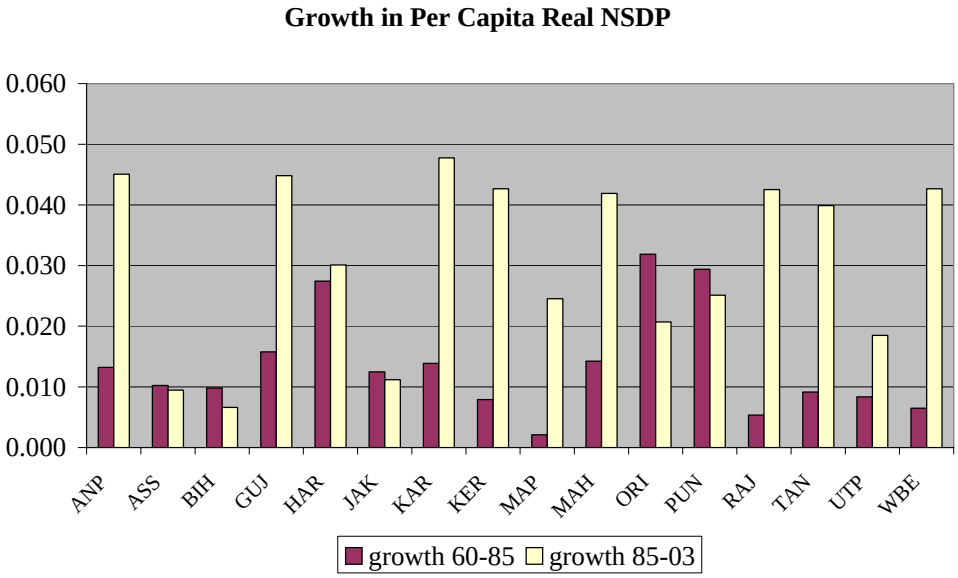
Figure 4. Growth of Log Real Output Per Capita in the 16 Major Indian States



We shall show later in this paper that there is some doubt about the precise timing of this break. Indeed, since the break has only been identified ex-post any attempt to identify the timing precisely is in any case subject to severe data-mining critiques. Nonetheless if we take a sufficiently long-term perspective there does seem to be good reason to place the breakpoint at some point in the 1980s. Without prejudice to subsequent discussion about precise timing, it is revealing to compare state-wise per capita growth rates before and after 1985, as shown in Figure 5.[7]

[7] A simple way to identify such a shift is by inclusion of dummy variables to capture a shift at a particular date. If we run a simple panel regression of growth rates on a

Figure 5



The chart displays very clear dividing lines, both across time and across states, which are most revealing if expressed in terms of convergence towards the global frontier. As background data US per capita GDP grew at 2.5% over the whole period, with only small differences between growth in the first and second sub-periods (2.6% vs 2.3%).[8]

Against this benchmark, which we can use as a reasonable proxy for the global frontier economy, only three Indian states, Haryana, Punjab and Orissa, showed any tendency to even marginal convergence in the first sub-period: they would be better described as just holding their own.[9]

The remaining states were all growing less rapidly than the frontier - indeed some, like Madhya Pradesh, were barely growing at all - so that almost all Indian states were to a greater or lesser extent, on the downward-sloping part of the "V" shown in Figure 1.

For the majority of states the contrast in the second period could hardly be any more striking. Eight states (Andhra Pradesh, Gujarat, Karnataka,

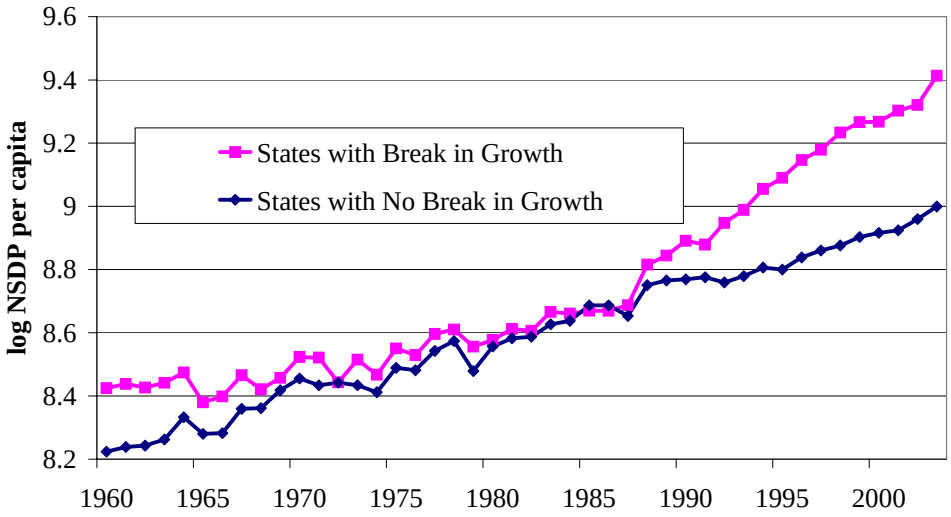
constant, and add a dummy variable that shifts from zero to unity in 1985, this is strongly significant on standard criteria. These are however clearly invalid on a data-mining based critique since we have chosen the date to maximise its notional signficance.

[8] Source: Penn World Tables

[9] Of these three states, Figure 2a shows that one, Orissa, had shown extremely rapid growth during the 1960s, but then had ceased any tendency to converge. These results are consistent with Datt and Ravallion (2002).

Kerala, Maharashtra, Rajasthan, Tamil Nadu and West Bengal) had per capita growth rates in the neighborhood of 4%, and were thus unambiguously converging; a ninth, Madhya Pradesh, managed a very significant shift in growth, but by only enough to roughly hold its own relative to the US. In the remaining states growth remained at a fairly similar rate to that in the previous sub-period. Within this group two states, Haryana and Punjab more or less maintained their relative position; but the remaining 5 states, Assam, Bihar, Jammu & Kashmir, Orissa and Uttar Pradesh, continued to lose ground.

Figure 6. Average Log Real Output Per Capita



Some recent research (e.g., Datt and Ravallion (2002)) has suggested that it has been predominantly rich states that have benefited from the growth turnaround. Figure 6 demonstrates in a simple way that this is not the explanation of the growth patterns shown in Figure 5. It compares the average per capita income of the nine states listed above which did significantly increase their growth rates with the average figure for the 7 states that did not. Given the way in which the two groups have been selected, it should be no surprise that the former group displays a clearly kink-shaped growth pattern, while the latter does not. But what is much more striking is that in the early 1980s, when the growth turnaround appears to have taken

place, the average income of both groups was almost identical.[10]

4 A first estimate of the "V-Factor"

A revealing way of capturing the different growth patterns discussed in the previous section is by a simple application of principal components analysis. Following Bai (2004) and Bai and Ng (2002; 2004), we assume that longer-term trends in output can be captured by a relatively small number of common factors that determine permanent (i.e, unit root) movements, i.e, a representation of the form

$$y_{it} = \beta_{i0} + \beta_{i1}F_{1t} + ... + \beta_{ik}F_{kt} + u_{it}; i = 1..N \quad (1)$$

$$\Delta F_{jt} = a(L)\varepsilon_{jt}^{P}; \quad j = 1..k \quad (2)$$

$$u_{it} = b(L)\varepsilon_{it}^{T}; \quad i = 1..N \quad (3)$$

where y_{it} is log output in state i; the F_{jt} are common factors that are subject to permanent shocks, the ε_{jt}^{P}; with the u_{it} capturing the remaining transitory dynamics.[11] The state level transitory shocks, the ε_{it}^{T} may in principle be mutually correlated. The β_{ij} are factor loadings on the common permanent factors. $a(L)$ and $b(L)$ are assumed to be stationary polynomials in the lag operator (defined such that for any variable x_t $Lx_t = x_{t-1}$) of the form $a(L) = a_0 + a_1L + a_2L^2 + ...$ Bai (2004) shows that as long as the u_{it} are stationary, consistent estimates of the common factors, and of the factor loadings, can be derived from the application of static principal components analysis.[12]

Figures 7 and 8 summarize the results of applying this approach to the 16 per capita state output series. There appears to be quite strong evidence that just two common permanent factors are sufficient.

[10] Precise figures are given in Table 1 below.

[11] An alternative interpretation with an identical representation writes output for each state as the sum of a permanent and a transitory component, where the innovations to the state-wise permanent component are correlated across the states.

[12] Bai and Ng (2004) outline an alternative approach which is consistent even when idioysncratic components are non-stationary. In this approach principal components analysis is applied to first differenced data, and the resulting components are cumulated. When this approach is applied to our dataset preliminary testing points to a larger number of factors. These are again strikingly V-shaped, and strongly correlated in terms of longer-term movements; but the larger number of factors makes interpretation more problematic. We shall provide a more detailed comparison of this approach, and statistical tests of the number of common factors, in a later draft of this paper.

Figure 7. First Two Principal Components of Statewise Log Real Output Per Capita

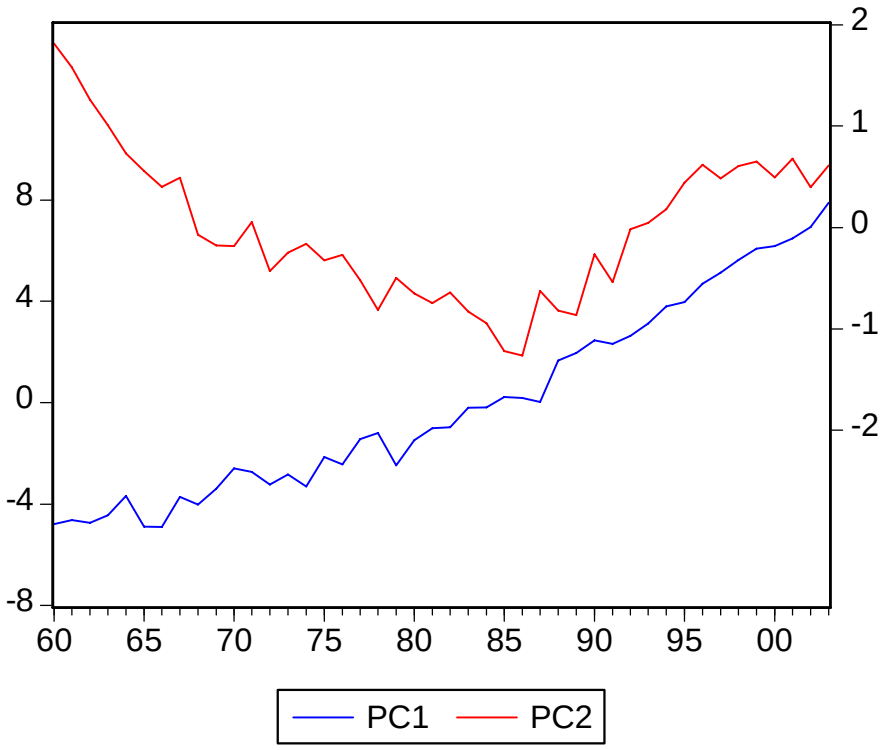


Figure 7 shows the first two principal components. The first captures long-term common growth trends (indeed it looks extremely similar to the cross-sectional average level of output). The second component, which has a strikingly similar shape to that shown in Figure 1, provides our first estimate of a common "V-factor".

If we assume that these two factors alone provide a sufficient representation of the common permanent components in state output, we can construct estimates of transitory components in each state defined by

$$\widehat{u}_{it} = y_{it} - \left(\widehat{\beta}_{i0} + \widehat{\beta}_{i1}\widehat{F}_{1t} + \widehat{\beta}_{i2}\widehat{F}_{2t}\right). \tag{4}$$

We show in the Appendix that for most, if not all states, the resulting series appear stationary. This is confirmed by formal tests which reject the unit root hypothesis both (strongly) for the panel as a whole, and also for each series in isolation.

A striking feature of the two principal components shown in Figure 7 is that the low point of the second component, the V-Factor, appears to correspond fairly closely to an apparent kink in the first principal component. The two components are, by construction, mutually orthogonal. But if we are prepared to admit the possibility that the permanent innovations in (2) may be correlated, we can straightforwardly specify the two factors in (1) such that the "V-Factor" alone provides a sufficient explanation of the differential shifts in growth performance summarized in the previous section.

Figure 8. Rotated Principal Components of Statewise Real Log Output per Capita

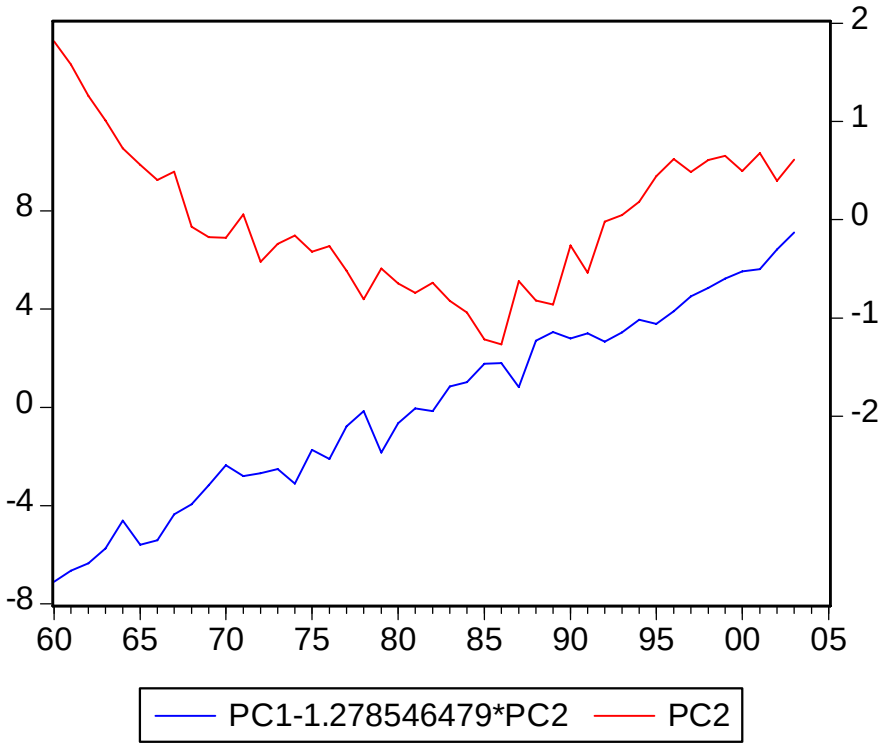


Figure 8 shows a simple rotation of the two factors, which sets $F_{1t} = PC_{1t} - \alpha PC_{2t}$; $F_{2t} = PC_{2t}$, where the coefficient α is defined such that the average growth of F_{1t} in the samples 1960-1985 and 1985-2003 is identical. With this specification of the two factors, the first factor explains long-term growth in individual states, while the V-factor ($\equiv F_{2t}$) alone explains all long-run shifts in growth.

Figure 9

V-Factor Loadings

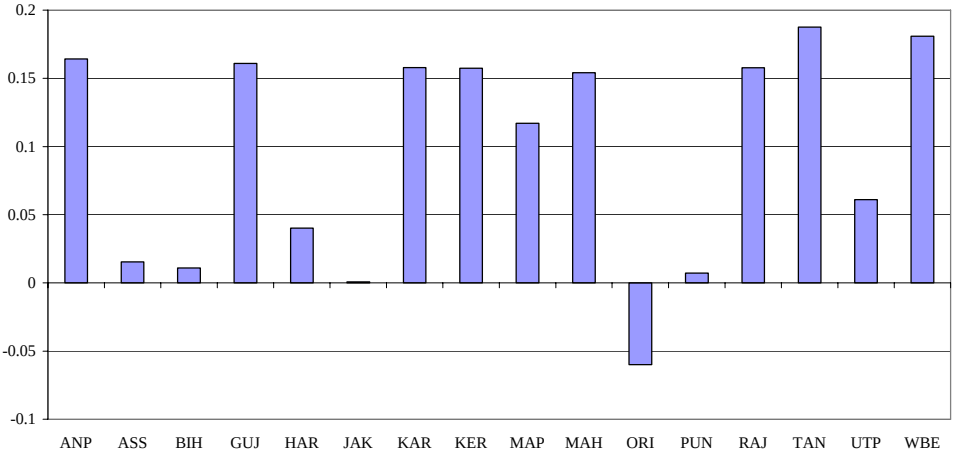


Figure 9 shows the implied factor loadings on the V-Factor (i.e., the $\widehat{\beta}_{i2t}$ in equation (1)). The set of states with strongly positive factor loadings exactly matches the set of states with clear shifts in growth rates shown in Figure 5. The remaining states have factor loadings on the V-Factor close to zero, or even, in the case of one state, Orissa, below zero. Since the majority of states have strongly positive V-Factor loadings the V-Factor also explains the shift in average growth.

It is perhaps worth clarifying that this representation, with two common permanent factors, with differential factor loadings for each state, is consistent with the evidence against convergence summarised in Figure 2. To see this consider the representation of any arbitrary pair of states, i and j, given the joint representation in (1) to (3), which implies

$$y_{it} - y_{jt} = \beta_{i0} - \beta_{j0} + u_{it} - u_{jt} + (\beta_{i1} - \beta_{i1}) F_{1t} + (\beta_{i2} - \beta_{i2}) F_{2t}. \qquad (5)$$

The sum of the the first four terms will be a stationary process given the assumption (supported by the data in our panel) that u_{it} and u_{jt} are both stationary. But given that the two common factors are non-stationary, we would would only observe even conditional convergence if $\beta_{i1} = \beta_{j1}$; $\beta_{i2} = \beta_{j2}$, $\forall i$, $\forall j$; while unconditional convergence would additionally require $\beta_{i0} = \beta_{j0}$; $\forall i$, $\forall j$. In broad terms this suggests that long-term growth

performance in the Indian states has (as we might expect) been hit by a sequence of common permanent shocks. What is perhaps more puzzling is that these shocks have had differential, and *sustained* differential effects in different states. In Section 8 we attempt to provide some economic rationale for this empirically driven representation.

Our two factor representation thus summarizes in compact form both the relative and absolute growth performance of total state output. It does so without making any assumptions about the date of any turnaround in growth, which simply emerges from the data.[13] We shall now go on to show that this is a repeating pattern in a much wider range of output data.

5 The V-Factor in Sectoral Data

We have thus far focussed solely on total output; but it is also interesting to examine the same factor decomposition at a sectoral level. Figure 10 compares the first two principal components of state-wise total output (15 series in total[14]) with those derived from state-wise sectoral output (13 sectors in 15 states, hence 195 series in total). The results are strikingly similar.[15] In one respect this is unsurprising, since if there is a common factor representation on a sectoral basis, it must aggregate up (at least to a log-linear approximation). But there is no guarantee that the factors that dominate long-term movements in sectoral output will dominate the aggregate, since in principle there might be significant sectoral factors that cancel out in the aggregate (e.g., if one sector were to systematically grow at the expense of another across all states). However it appears that the growth factor and the V-factor dominate at both the aggregate and sectoral level.[16]

[13] Note that, while the rotation of the two factors shown in Figure 7 does, for convenience, assume a break-point in 1985, this only affects estimated factor loadings. It makes no difference to the estimate of the V-Factor itself, nor to the total contribution of the two factors.

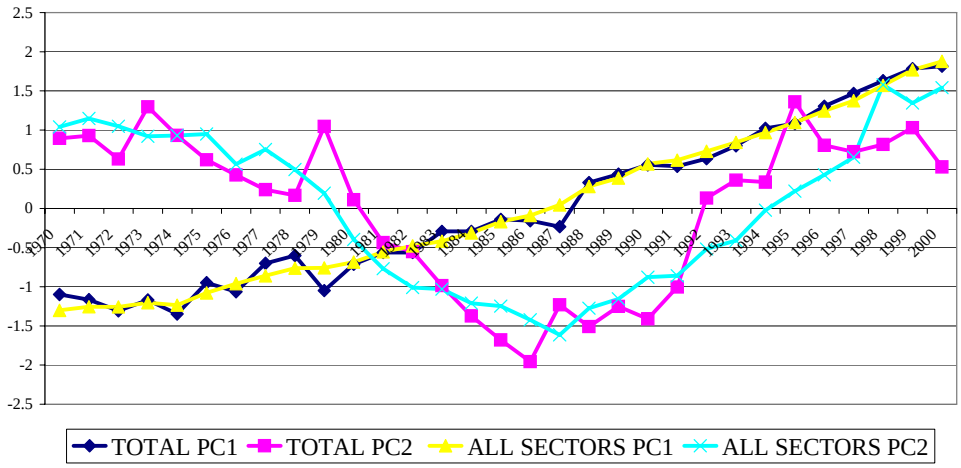
[14] We do not currently have sectoral data for Jammu & Kashmir; we do however aim to add these in a future draft for consistency with total data.

[15] Note that due to current data limitations we carry out sectoral analysis on data from 1970-2000.

[16] This is borne out by unit root tests on sectoral output after stripping out the effect of the two factors from state-wise sectoral output levels, which indicate stationarity in the great majority of cases.

Figure 10

First 2 Principal Components, Total Output vs All Sectors, 1970-2000



This result is in some respects all the more striking because, while the nature of the two factor representation remains clearly visible, its impact is by no means uniform across sectors. Thus while some sectors appear in general to have more significant impacts of the V-Factor, there appear to be important V-Sectors even in states where the aggregate effect of the V-Factor is close to zero (and, to a lesser extent, vice versa).

Given the multi-dimensional nature of the analysis, the detail of the factor-based sectoral analysis is relegated to the Appendix. A few charts do however provide some summary insights.

First, as background, it is helpful to be reminded of some straightforward sectoral growth accounting stylized facts. Figures 11 and 12 show the contributions to total aggregate growth of five broad sectors over the two fifteen year periods, 1970-1985 and 1985-2000.[17] Since the two periods are of the same length the two charts are directly comparable, and hence are

[17] If $y_{y,t}$ is sectoral output, the contributions in Figure 10 are caculated as

$$cont_j = \frac{(y_{j,1985} - y_{j,1970})}{Total_{1970}} \times 100$$

which sums to the percentage change in the total if it is precisely equal to the sum of components in both years. Due to data inconsistencies and shifts in base years the identity does not hold precisely; however the chart shows that the discrepancies are fairly minor. Private services are defined as the sum of transport & communication, trade, banking and insurance and other services.

shown on the same scale. While output per capita grew by only between zero and 40% in the first sub-period, in the second sub-period it roughly doubled in the majority of states.

Figure 11

Sectoral Contributions to Growth, 1970-1985*

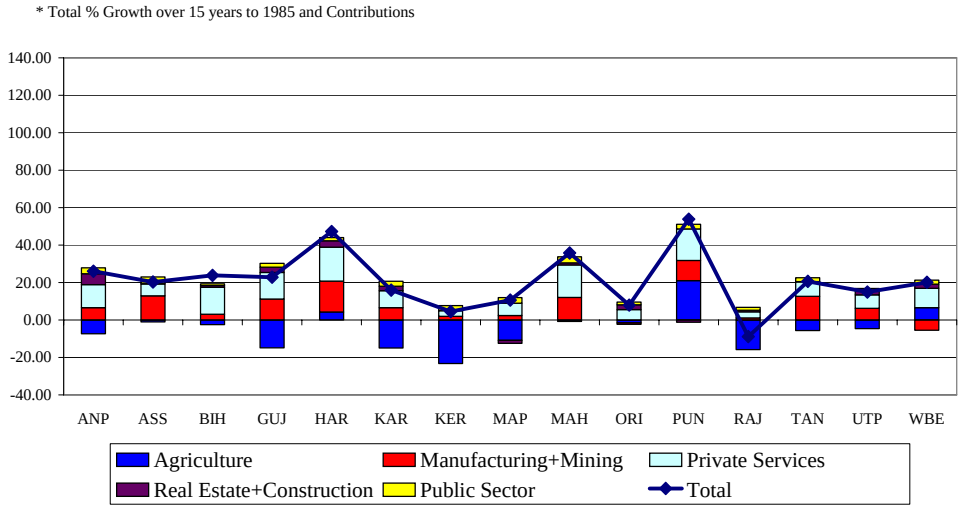
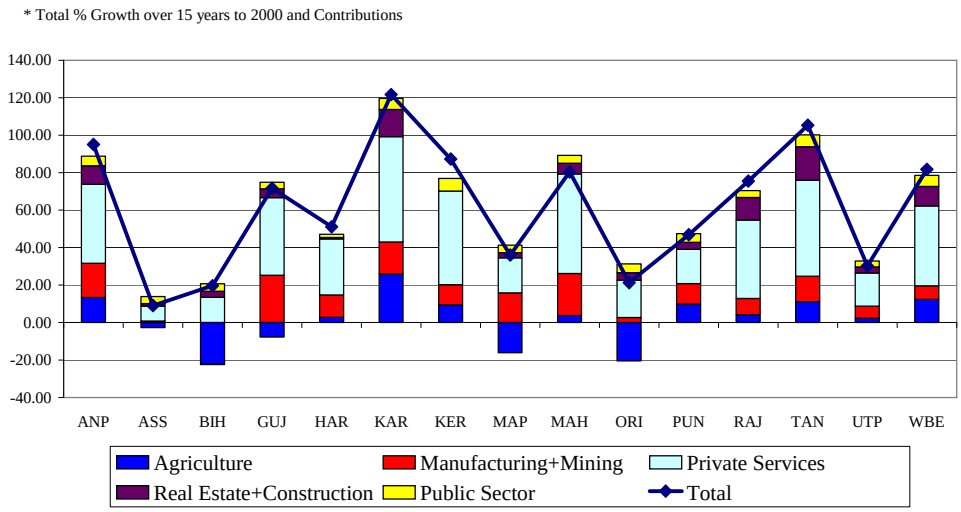


Figure 12

Sectoral Contributions to Growth, 1985-2000*



Key features illustrated in the two charts are:

- The growth up-turn in the latter period was, as is quite well-known dominated by up-turns in growth in the (tertiary) service sector.[18] But even in the earlier period this sector was (at least relatively speaking) already making an important contribution

- There were also significant improvements across all sectors in most states.

- While agriculture remains the dominant element in total output in most states, it made at best only small contributions to growth in either period. However, the latter period did at least see distinctly fewer negative contributions from agriculture.

Figure 13

Sectoral V-Factor Loadings

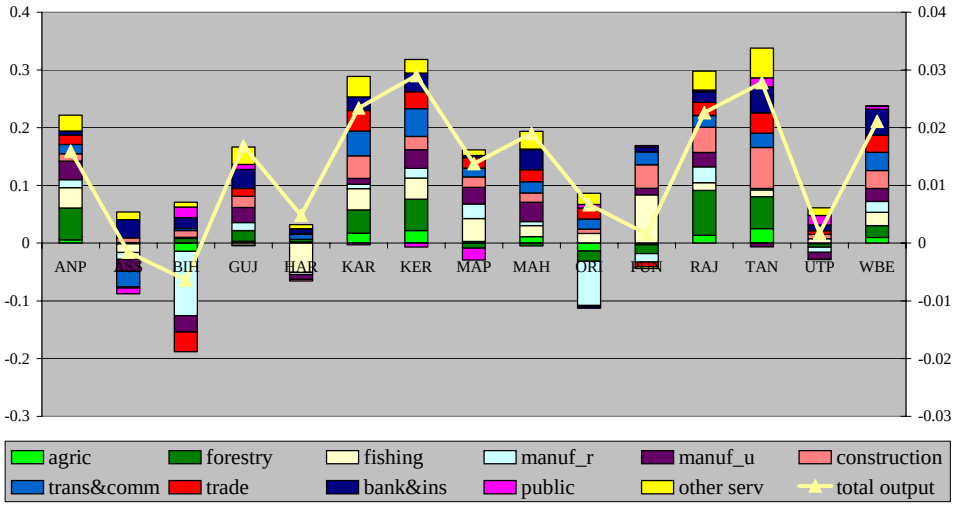


Figure 13 relates these features of the data to the factor representation. It compares the sum of the factor loadings on the V-Factor for all 13 industrial sectors in each state with those for output as whole. At the aggregate level, states with high V-Factor loadings had more marked turnarounds in growth. The sectoral counterpart to this is that in these same states the V-Factor affected more sectors, and with a typically higher weight.[19] The

[18] These results are consistent with Balakrishnan and Parameswaran (2007).

[19] The total V-Factor loading is, up to a log linear approximation, a weighted average of the sectoral V-Factor loadings, whereas the chart in effect compares it with the unweighted average; but this makes relatively little difference to the results.

impact of the V-Factor was typically spread across a wide range of sectors. The impact was however by no means uniform, whether across states or across sectors.[20]

6 Dating the growth turnaround

In the preceding two sections we have established that we can derive a satisfactory representation of state level per capita output, both in aggregate, and at the sectoral level, in terms of just two factors: a growth factor and a V-Factor. These can be constructed in such a way that the growth factor has a roughly constant impact on output over time, while the V-Factor alone explains the turnaround in growth.

Given this representation, an obvious question is when the turnaround in growth actually began. As noted in Section 2, some past research has suggested, on the basis of analysis of total output that there was a break in growth early in the 1980s, or possibly even as early as the late 1970s. This conclusion does not appear out of line with Figure 1, which compares All-Indian output per capita with the same series for the United States. As such, this result appears somewhat surprising, given the lack of any obvious switch in policy that might have brought this change about.

Our own more disaggregated factor-based analysis suggests a rather different conclusion. The estimates of the V-Factor shown in Figures 7 and 10 have low points in the mid- to late 1980s - much more consistent with what we know about the history of policy.

It would be tempting to claim that this more plausible dating of the growth turnaround is due solely to superior statistical techniques. It is indeed certainly the case that our analysis does draw on a much wider range of data. The timing of the turnaround in the mid-1980s also appears quite robust on our dataset. If, for example, we partition the dataset in different ways before deriving principal components, by using data for individual sectors in all states, or for all sectors in a given state, in the great majority of such partitioned datasets we also find a V-shaped 2nd principal component, with its low point in the mid- to late 1980s.

However, we would not wish to over-sell this conclusion, since an important part of the explanation of the differences between our conclusions and

[20] See Appendix for more detailed analysis of the sectoral factor loadings.

those based on aggregate data can be ascribed to data discrepancies.

While state-wise output estimates are constructed using the same conventions as the total figure, there are non-trivial discrepancies between the All-India figure and the sum of the state-wise output figures. These discrepancies, which can be ascribed only partly to the output of supra-state bodies,[21] are typically reasonably stable; however, the 1980s appears to have been a period in which the discrepancy showed more significant movements. These make inferences about precise timing of the turnaround fraught with difficulties.

Figure 14. Alternative Estimates of Average Log Real Output Per Capita

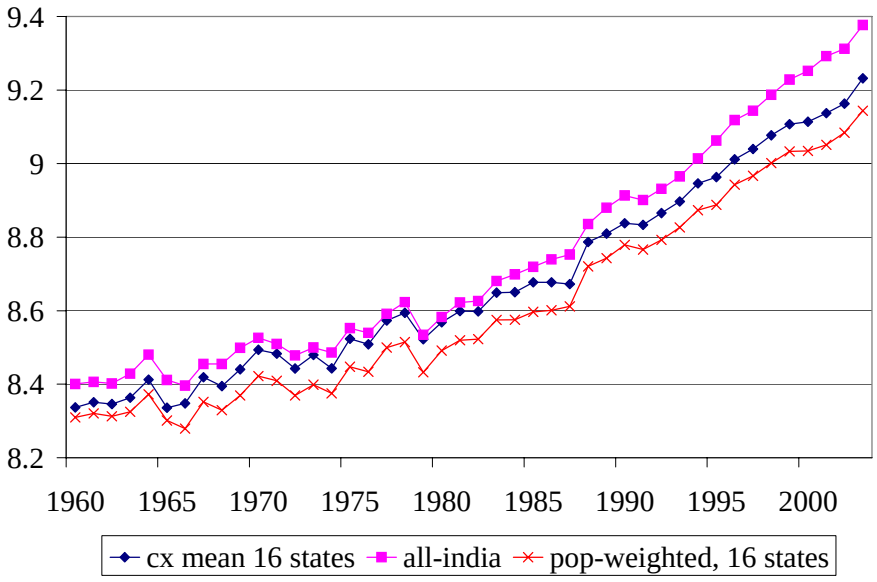


Figure 14 illustrates the nature of the problem. It compares the published figure for log per capita output with equivalent series derived from state-wise figures, both unweighted and population-weighted. The latter figure is clearly the appropriate figure to compare with the All-India total, and should in principle provide a good match, given that the 16 states in our dataset include 97% of the population. However, the unweighted average is what is relevant to our principal components analysis, which treats all states symmetrically.

[21] The items not covered in the NSDP numbers are (i) defence and other para-military forces, (ii) government offices abroad, (iii) foreign offices of LIC and GIC, (iv) Bombay off-shore activities, (v) deep sea fishing and, (vi) net income earned from abroad.

The figures typically move in-step with each other, albeit with non-trivial differences in levels. However the chart shows that in the 1980s in particular the discrepancies widened in a way that unfortunately clouds the issue of the timing of the growth turnaround. All three series fell sharply in 1980. The all-India figure then grew fairly steadily thereafter, while the unweighted average of the state-wise figures levelled off during the mid-1980s before picking up again sharply towards. To a reasonable approximation, analysis of aggregate figures dates the growth turnaround at or near the local minimum in 1980, while our disaggregated analysis dates it at a local minimum in the mid-1980s which simply does not appear in the aggregate data.[22]

We would argue that some degree of uncertainty about precise timing is endemic to this type of analysis. Short-term volatility of growth rates is very high in comparison to the magnitude of any plausible longer-term shifts in growth, such that any signal is easily dominated by short-term noise. However, over the longer term the reverse is the case, since long-term output levels - which feed into our principal components analysis - are dominated by average growth rates. The data thus provide much stronger evidence for the existence of a long-term V-factor than they do for the precise location of its low-point.

7 Correlates of the V-Factor?

Since we have established that our V-Factor can capture the major features of the Indian growth turnaround both in aggregate and at a disaggregate level, it would obviously be of considerable interest if we could establish some causal factor, movements in which were correlated with the V-Factor.

We do not claim to have found any such causal factor. Indeed it would be surprising if we had.[23] Nonetheless, we can at least provide some preliminary evidence of what correlates with, and (perhaps of equal interest) what does not correlate with the V-Factor.

Inevitably, if frustratingly, evidence that we have on state-wise indica-

[22] Note that the series in Figure 1 is also complicated by short-term volatility of US output.

[23] As a comparision, consider the amount of research effort that has been devoted to analysis of, for example, productivity movements in the United States alone, without providing a clear-cut explanation in terms of causal factors.

tors that we might hope to relate to the V-Factor is, more often than not, only infrequently sampled - most notably on a decadal basis in census years. For such series we simply do not have enough of a time series dimension to enable a comparison with our estimates of the V-Factor. We can however at least ask if the values of such indicators immediately before the low-point of the V-Factor would have given some indication of which states were likely to benefit differentially from the turnaround. If nothing else this allows us to dispense of some candidate explanations.

Table 1. V- vs Non-V States: Values in 1981

	NSDP per capita, 1993 prices	Population, millions	Population growth rate, 1971-1981	Kilowatt/ Hours per capita	Fixed Investment % of NSDP	Development Spending, % of NSDP	Literacy Rate	% Urban Population	Agriculture % of NSDP	Manufacturing % of NSDP	Registered Manufacturing % of NSDP
V-States	5337	44.7	2.2	196.5	27.9	10.7	49.0	26.4	39.6	16.9	9.9
Non-*V* States	5497	37.3	2.2	187.2	31.0	12.7	38.2	17.7	47.9	10.3	5.7

Table 1 shows values of a range of state-wise indicators in the census year immediately preceding the turnaround. To summarize the data we split state into the same groups as in Figure 6, which, given the V-factor loadings in Figure 9 we can loosely characterize as "V- vs Non-V" states.[24]

If we focus initially just on the first five columns of Table 1, the most striking aspect of the data is just how *little* difference there was between V- and Non-V states in terms of some important indicators, just before the growth turnaround. As already noted in relation to Figure 6, just before the apex of the V-Factor the two groups had virtually identical incomes per capita. They also had fairly similar populations and virtually identical population growth rates and investment rates, hence explanations based on differences in neo-classical growth model parameters do not appear to apply.[25] As a proxy for supply constraints, initial electricity generating capacity was also very similar

Moving further across the table we do observe at least some differences:

- V-States were on average more urbanised and more literate;[26]

[24] The analysis of Table 1 could of course equally well be carried out without reference to the V-Factor per se. A future version of this paper will provide a more systematic comparison of state-wise values of indicators such as those shown in Table 1 with state-wise V-factor loadings.

[25] Although we should note the important caveat that statewise investment figures are highly volatile, and are known to be prone to very significant measurement errors.

[26] The differences in literacy are inevitably accentuated by the inclusion of Kerala (with

- They were somewhat more industrialised, and somewhat less dependent on agriculture;
- They spent somewhat *less* on development spending (revenue expenditure) than non-V states.

These figures are of course simply a snapshot, and therefore may not tell us much about subsequent behaviour of individual states.

In the case of some indicators initial differences did not correspond to any subsequent correlation with the V-Factor, and thus seem unlikely to provide any clues to causal relationships. A specific example of this type is the share of registered manufacturing in total output. Table 1 shows that, in line with the analysis of Rodrik & Subramanian (2005) a strong subsequent growth performance of a given state appears to have been related to a high initial share of registered manufacturing. But the detailed sectoral analysis in the Appendix shows that registered manufacturing was not typically a "V-Sector": i.e., while it enjoyed reasonably high average growth, it did not experience any obvious turnaround in growth. Thus a relatively high initial share of registered manufacturing appears to have been at best a catalyst for better growth performance, not a deep causal factor.[27]

Some patterns do however appear to have had more long-standing effects. Figures 13 and A3 (see Appendix) show that agriculture typically had quite low loadings on the V-Factor. Thus states with relatively high initial shares of agriculture were in a less good position to benefit from the impact of the V-Factor. On the other hand the same charts show that private service sectors typically had higher V-Factor loadings. While V- and Non-V States had very similar initial shares of private services,[28] we can at least guess that their relatively higher levels of literacy and urbanisation shown in Table 1 may have given them a comparative advantage in capturing the benefits of the V-Factor in private services.

As already noted, apart from the sectoral output series we have relatively few true time series on a statewise basis from which we can hope to extract

a literacy rate of 82% in 1981) in the V-States; however even excluding Kerala the average literacy rate for the V-States was 45%.

[27] This conclusion is indeed very close to the argument originally used by Rodrik and Subramanian, who rationalised the apparent signficance of the initial share of registered manufacturing as symptomatic of an initial change in the emphasis of central government policy in favour of relatively large-scale business.

[28] Defined as in footnote 17.

further clues. We can however point to two fairly striking features of the data.

The first is the role of the public sector. As already noted, V-states were initially spending less (albeit only marginally on development spending). But this differential pattern also has its counterpart in subsequent developments. Figures 12 and A3 show that the recorded output of the public sector (which is largely driven by public sector employment) had an average V-Factor loading close to zero. Perhaps more strikingly Figure A3 shows that state-wise V-factor loadings for the public sector were actually inversely correlated with those for total output. Thus, V-states have tended to decrease public spending since the 1980s, and non-V states have tended to increase it. A very similar pattern is evident in development spending. This is intuitive since following Ozler, Datt, and Ravallion (1996), we measure development expenditures by revenue expenditures. Since the mid 1980's, many V-states (such as Gujarat, Maharashtra, and Tamil Nadu) have decreased public spending on manpower (public employment) and increased capital investments. This suggest that revenue expenditures are a symptom of poverty, as supply constrained economies reduce capital investments and increase revenue expenditures.[29] Related to this, most infrastructure projects require substantial fixed costs. Such projects cannot be undertaken unless a region has sufficient absorptive capacities and income is higher than a critical threshold level. States with higher levels of income also have more fiscal revenue to pay for more ambitious infrastructure projects.

A second striking correlation with the V-Factor shows up in one very crucial aspect of the supply side of the Indian economy, namely electricity generation. Table 1 showed that V-States had no significant initial advantage in this respect; however the subsequent profile of this series shows a distinct pattern across the two groups.

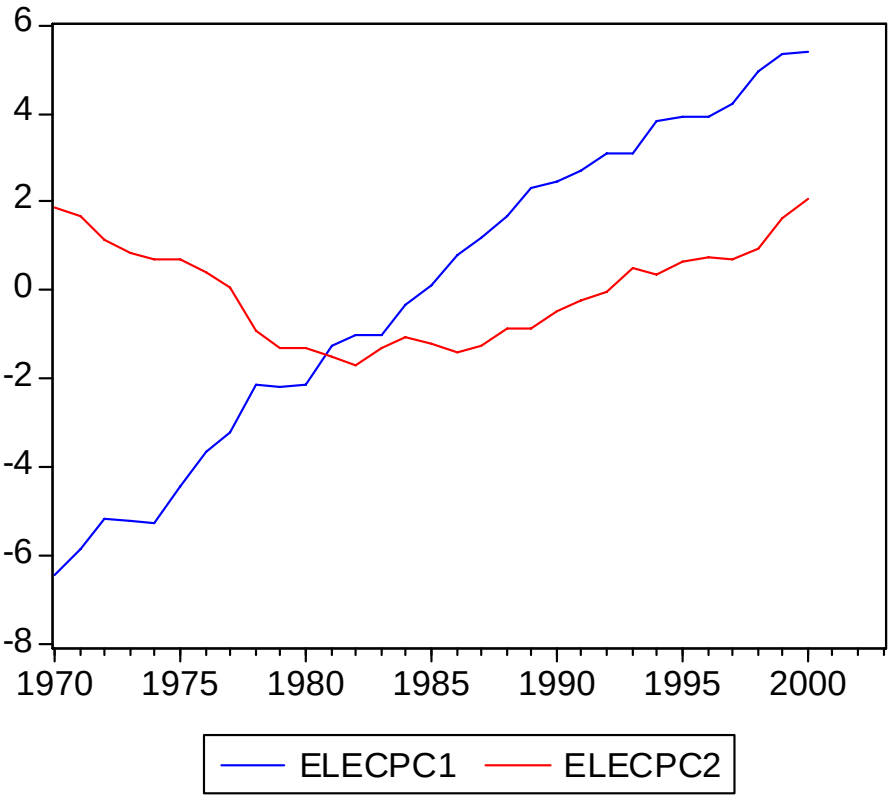
Figure 15 shows that the first two principal components of per capita electricity generation display the familiar pattern we have seen in output data, with the second component providing yet another V-Factor. The signs of factor loadings of individual states on the V-Factor for electricity match extremely well (in 15 out of 16 states) with those of total state output on the output V-factor.[30] In a rich country this correlation would not be

[29] See Ghate (2008) who constructs a median voter model of infrastructure investments and regional divergence in the Indian context.

[30] A caveat relating to this series in more recent years is that electricity transmis-

especially interesting: indeed in some research on the US economy electricity generation has been used as a short-term proxy for output itself. In the case of India, where electricity supply interruptions are still common, the correlation is less obviously trivial. Electricity supply is for some activities close to being a binding supply constraint. The correlation between the two V-Factors, and their factor loadings shows that V-states have typically been better at progressively releasing themselves from this constraint.[31]

Figure 15. First Two Principal Components of Per Capita Electricity Generation



sion across states has also improved across states with an increase in inter-state flows of electricity affecting irrigation and agriculture positively. Thus statewise production of electricity may over time become a less good proxy for statewise consumption, which is what we would ideally like to measure, but for which we have much shorter and incomplete samples of data.

[31] Nagaraj, Varoudokis and Venganzones (1998) also find evidence that per capita total consumption of electricity, and per capita industrial consumption affect state economic performance positively. However they conclude that the percentage of villages electrified or other physical infrastructure variables such as length of the railway network has no clearly identifiable positive impact. We find a similar lack of correlation in our own dataset.

8 Trying to make sense of the V-Factor

Can we rationalise the existence and ubiquity of the V-Factor, and the limited evidence we have of its correlates, with any underlying economic model? We can get some insights into this question from considering a fairly general model of convergence of the form

$$\Delta\left(y_{it+1}-y_{t+1}^{US}\right)=\alpha_i\left(y_t^{US}+s_{it}+s_t^{India}-y_{it}\right)+\Delta TFP_{it+1}-\Delta TFP_{t+1}^{US}+\varepsilon_{it+1} \tag{6}$$

where y_{it} is log output per capita for state i , the s_{it} and s_s^{India} variables captures factors that determine steady-state output relative to the frontier represented by y_t^{US}, log output per capita in the United States, for individual states and for India as a whole; TFP_{it} and TFP_t^{US} is growth rate of total factor productivity in state i and in the United States and ε_{it} captures short-run cyclical factors.

As noted in relation to the discussion of Figure 5, to converge towards US per capita output required growth greater than around 2% per annum. Before the mid-1980s very few Indian states achieved this, and if so only marginally. In contrast, since the mid-1980s the V-states have all been converging (though at very different rates); while non-V states have been barely converging, or have continuing to diverge.

The simple framework of (6) offers a range of possible ways of accounting for the all-India pattern; but not all such explanations are so readily applicable to the relative performance of different states.

It seems reasonable to argue that the sum of the last three terms on the right-hand side of (6) is unlikely to provide an adequate explanation of longer-term trends. In standard Cobb-Douglas type technology models TFP growth shocks are common across all economies and hence cancel out precisely. But even if they are country specific, such relative shocks might reasonably be assumed to have a stationary distribution. The same applies to the short-term error term, ε_{it+1}. Thus we need to look for an explanation somewhere in the first term.

One possible (and rather pessimistic) interpretation of the earlier period was that the bracketed "convergence" term (the term multiplied by α_i) was on average close to zero - ie, that most, or possibly all Indian states were, conditional upon the s_{it} and s_t^{India} processes, fairly close to their steady-state

values. The downward drift in most states' relative output levels would, according to this interpretation, be interpreted either as a succession of bad relative TFP growth shocks, or possibly (and even more pessimistically) as a downward drift in s_t^{India}.

It is harder to continue the logic of this explanation after the growth turnaround. One obvious candidate explanation is that at some point in the mid-1980s there was a shift in steady state output levels due to successive liberalisations driven by the centre. Many of these shifts were manifestly common across all states, hence it is reasonable to attribute them to changes in the common Indian steady state factor s_t^{India}. Given the subsequent doubling of growth rates (and an even more dramatic change in rates of convergence), then, conditional upon a reasonable degree of stability in the other elements on the right-hand side of (6), including rates of convergence, the implied changes in s_t^{India} must have been quite dramatic. Rodrik and Subramanian (2005) argue that this is plausible because India was well away from its production possibility frontier.

But since these changes were common across states, the great puzzle presented by the differential impact of the V-Factor is why any such shift in s_t^{India} did not have largely symmetric effects across the states. If we are to pursue this line of explanation, we have to look for equivalently, or even larger shifts in the state-specific s_{it} factors that determine steady-state output, ocurring more or less contemporaneously with the India-wide shifts. It is hard to rationalise such dramatic shifts either on the basis of what we know about differential statewise policy changes, or on the basis of the very limited evidence we have presented of state-wise correlates with the V-Factor.

But there is an alternative explanation of the same pattern which seems somewhat easier to reconcile with both the all-India and statewise evidence. The analysis of these shifts has implicitly assumed that the state-specific rates of convergence, α_i were both strictly positive and reasonably similar across states. But an alternative explanation would attribute the pattern of the evidence largely to the α_i themselves. On this interpretation, and consistent with the arguments of Rodrik and Subramanian, the bracketed expression in the first term was *not* necessarily close to zero in the first period; but failure to converge to the global frontier was largely due to the α_i being so close to zero that differences between actual and steady state

income levels had essentially no impact. The turnaround in growth and its differential pattern would then be attributed to some combination of a common shift in s_t^{India} and statewise differences in the α_i. A differential impact of the all-India shock might be attributed to different values of α_i, with non-V states, by implication, having α_i values extremely close to zero, thus closing off any convergence response.

But a further possibility is that the differential impact of the V-factor reflects not just differential responses to common shocks to the steady states, but also shocks to the α_i themselves. One interpretation of convergence is as a process of arbitrage, driven by international differences in factor returns. Even in a frictionless model of convergence, low values of α_i can reflect low intertemporal elasticities of substitution, with the limiting case of $\alpha_i = 0$ corresponding to an elasticity of intertemporal substitution of precisely zero (Barro & Sala-Martin, 1992; Campbell, 1994). But models with frictions can also generate similar results, even when the true elasticity is positive. On this interpretation, the reforms of the 1980s and thereafter may not just have raised steady-state output levels, but may also have reduced frictions; with some states being better capable of exploiting the implied arbitrage opportunity. In this interpretation all states might in principle ultimately converge on very similar long-run output levels, but differential speeds of convergence would imply that they would appear, during the course of this process, to be systematically diverging, as Figure 2 suggests has been the case.

9 Conclusions

We have presented evidence of a common "V-Factor", derived from principal components of both total and sectoral output levels for the Indian states, that appears to capture well.long-term developments in both the absolute levels of output per capita and its cross-sectional distribution. The V-Factor appears to have its apex in the mid- to late 1980s, which is more consistent with the history of policy than previous studies, such as Rodrik and Subramanian (2005) that have dated the turnaround to the beginning of the 1980s. Factor loadings on the V-Factor allow us to identify V- and non-V states and industries; we have also presented some preliminary evidence of correlates of the V-Factor. The differential performance of V- vs non-V

states presents a puzzle to standard models of convergence if all states are assumed to converge at roughly the same rate, but is somewhat easier to explain if convergence rates differ.

Appendix

A Description of the Dataset

We utilize state level data from various state economic surveys, the Reserve Bank of India, the Census, and CSO publications. We incorporate data from the EPW Research Foundation (2005) dataset, the Ozler Datt and Ravallion (1996) dataset , the Center for Monitoring the Indian Economy (CMIE) dataset, and the Besley and Burgess (2000) dataset. A full listing of all the variables in the data set, their duration, and their sources is available from http://www.isid.ac.in/~cghate/ chetanresearch.html. We briefly describe the key variables.

NSDP Data: The NSDP data have been assembled from various tables in the EPW Research Foundation dataset. Our final dataset includes annual RNSDP (Real Net State Domestic Product) and PCRNSDP (Per Capita Real Net State Domestic Product) observations for 31 states from 1960 - 2005. The observations have been spliced so that all states have RNSDP figures in constant 1993-1994 prices. Our method of splicing ensures that our measures of state RNSDP are largely immunised from the impact of various changes in state definition.[32]

Population Data : The Population data has been tabluated from Census figures, with a common compound growth rate applied across decadal observations to impute annual observations for each state. We cross check these figures with population figures obtained by simple extrapolation: (NRSDP/PCNRSDP)* 10000000. Both the Census figures and extrapolated figures are consistent with each other. We also use the RURAL POPULATION and URBAN POPULATION proportions from various rounds of the NSS surveys to give us a full series of rural and urban annual population figures for 31 states from 1960 - 2005.

Sectoral Data : We report sectoral (primary, secondary, and tertiary) data for 16 major states (where available) using data from the EPW Research Foundation. All data have been spliced so that the sectoral data are in constant 1993-1994 prices. The variables are: Agriculture, Forestry and Logging, Fishing, Mining and Quarrying, Manufacturing (Registered

[32] These changes mainly affect Bihar and, to a lesser extent, Madya Pradhesh and Assam. Details of precise methodology are available from the authors.

and Unregistered), Construction, Electricity, Gas and Water Supply, Transport, Storage and Communication, Railways, Transport by other means and Storage, Communication, Trade, Hotels and Restaurants, Banking and Insurance, Real Estate, Public Administration, and Other Services

Development Expenditure Data : We combine data from the Ozler, Datt, and Ravallion (1996) dataset with various state statistical abstracts to get revenue expenditure shares of nominal state domestic product from 1960-2000. This includes state expenditures on Education, Sports, Art and Culture, Medical, Public Health & Family Welfare, Water Supply and Sanitation, Housing, Urban Development, Agriculture and Allied Activities, Rural Development, Irrigation & Flood Control, Energy, Industry and Minerals, Transport and Communication, Social Security, Welfare, and Roads and Bridges.

Political Variables: We report data on the number of registered trade unions, number of political parties in state governments, and state wide representation in the Lok Sabha. These data are from various issues of the Statistical Abstracts.

Rainfall Data **:** We report average monthly (from June to September) rainfall data, as well as the standard deviation of monthly rainfall data, at the state level for 16 major states from 1960 - 2000.

Infrastructure Variables: We report data on Commercially Consumed Electricity, Electricity Consumed by Agriculture, Gross Electricity Generation by Utilities, Gross Electricity Generation (Non Utilities), Railways density of route length per'000, sq km, Road density per '000, sq km, Motor Vehicles Density Per Sq Km of Geographical Area, Percent of unsurfaced roads to total roads, and Circle–wise telephone exchanges. These variables are for the 16 major states from 1970 - 2001.

Agriculture Land Usage: We report land usage data for agriculture based on various land holding sizes: areas operated by marginal, small, medium and large holdings, for 16 major states, from 1970 - 2003.

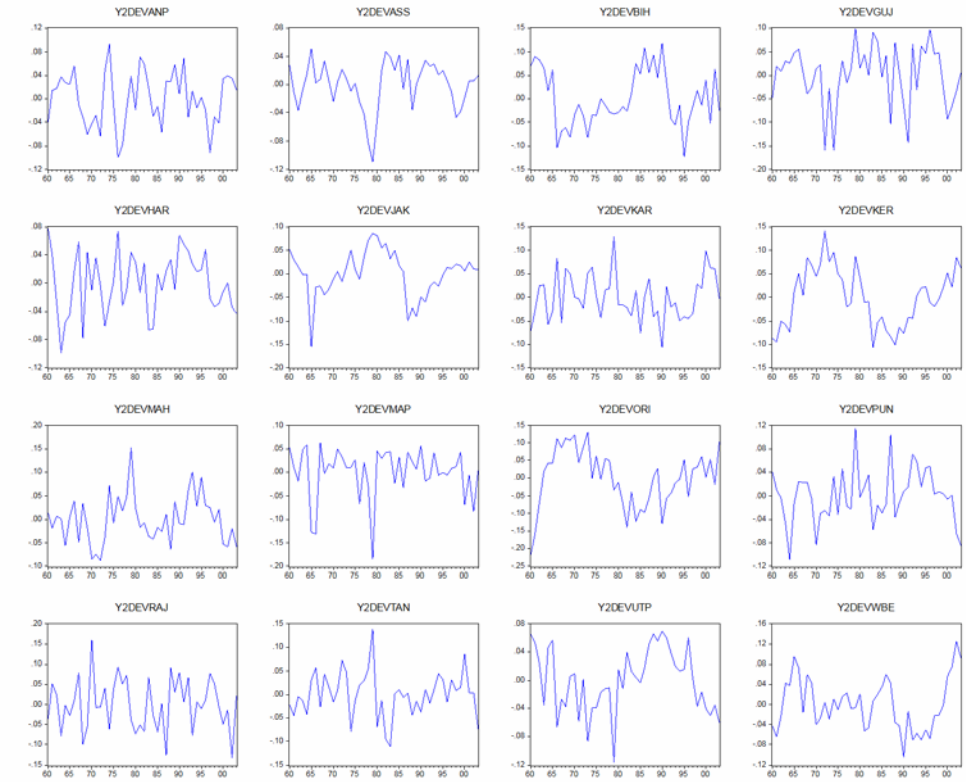
Public Finance Data : We report state level public finance data for 16 major states from 1980 onwards from various issues of the state economic surveys. These include the Gross Fiscal Deficit, Revenue Deficit , Primary Deficit, Own Tax Revenue, Own Non-Tax Revenue, Grants from the Center, States share in Center's Taxes, Revenue Receipts, Capital Receipts, Total Tax Revenues, and State Tax – Net State RGSDP ratios. These data are

from various issues of the state economic surveys.

B Transitory Components

Figure A1 shows estimated transitory components defined as in (4).

Figure A1. Estimated transitory components ($\widehat{u}_{it}$) in statewise log real output per capita



Tables A1 and A2 below show, respectively, ADF tests on individual statewise transitory components, and alternative panel unit root tests. Table A1 shows rejection of individual unit roots at below 10% probability levels in 14 out of 16 states. Panel unit roots strongly reject unit roots whether or not a common AR coefficient is assumed under the alternative.

Table A1, ADF tests

Series	t-Stat	Prob.	Lag
ANP	-4.47	0.00	0
ASS	-3.31	0.02	0
BIH	-2.39	0.15	1
GUJ	-5.36	0.00	0
HAR	-5.32	0.00	0
JAK	-3.07	0.04	0
KAR	-5.73	0.00	0
KER	-2.51	0.12	0
MAH	-4.29	0.00	0
MAP	-6.62	0.00	0
ORI	-2.84	0.06	1
PUN	-5.08	0.00	0
RAJ	-6.87	0.00	0
TAN	-5.19	0.00	0
UTP	-4.04	0.00	0
WBE	-2.76	0.07	0

Table A2, Panel Unit Root Tests

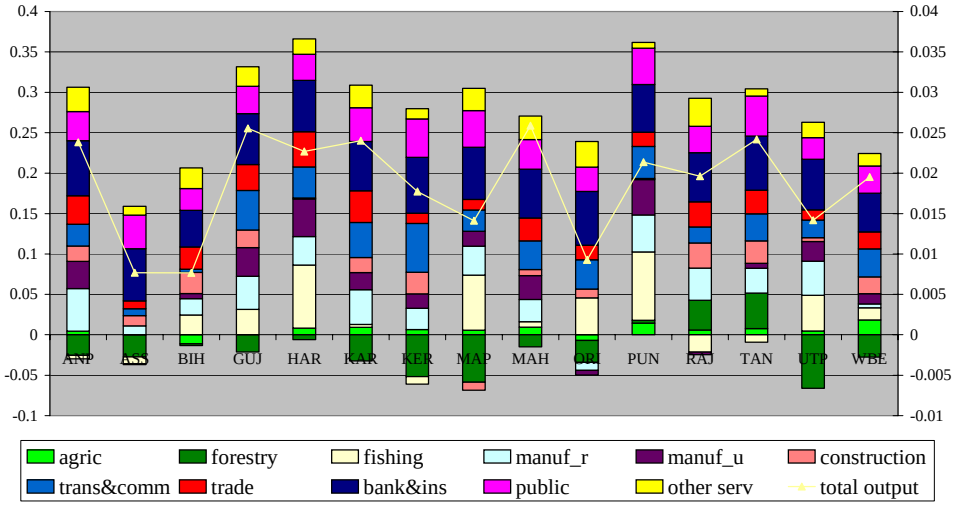
Method	Statistic	Prob.**
Null: Unit root (assumes common unit root process)		
Levin, Lin & Chu t*	-10.75	0.00
Null: Unit root (assumes individual unit root process		
Im, Pesaran and Shin W-sta	-12.95	0.00
ADF - Fisher Chi-square	225.91	0.00
PP - Fisher Chi-square	233.90	0.00
Null: No unit root (assumes common unit root proce		
Hadri Z-stat	-1.87	0.97

C Sectoral Factor Analysis

While the main paper has focussed primarily on analysing the properties of the V-Factor, it should be borne in mind that the greater part of the long-term variation in output, both at the aggregate and at the sectoral levels, is (by the very nature of principal components) attributable to the first factor, which we term the growth factor. Figure A2 compares sectoral growth factor loadings alongside the loadings for total output, across all 15 states.

Figure A2

Sectoral Growth Factor Loadings



With a few important exceptions most sectors have grown over the long-term, and thus most have positive growth factor loadings. But differential loadings mean that the growth factor has also had important impacts on sectoral distribution over time. The most marked exceptions to the general pattern of positive growth factor loadings are agriculture and forestry, for which loadings are typically close to zero or even negative At the other extreme banking & finance typically has a high growth factor loading (even in relatively slow-growing states) but appears less affected by the V-Factor.

Figure A3 shows an alternative comparison. It summarises sectoral properties in terms of two key aspects of the factor loadings on both the growth factor and the V-factor: the average loading of the sector across the states, and the state-wise correlation of factor loadings for a given sector with the pattern of statewise factor loadings for output as a whole.

The majority of sectors have positive loadings on both the growth factor and the V-Factor. But there are some interesting exceptions. Agriculture has on average low weightings on both. Registered manufacturing has a reasonably high weight on the growth factor, but a negative (albeit near-zero) weight on the V-Factor: i.e., it was a relatively fast growing sector on average; but in contrast to most sectors it did not typically show any marked pick-up in growth after the 1980s (this is also evident by looking at the raw data).

Figure A3

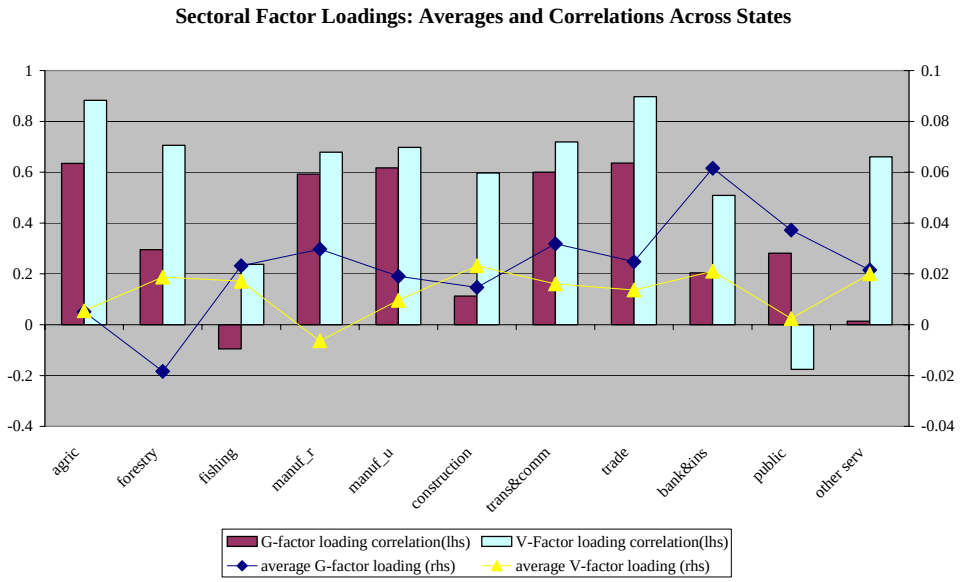


Figure A3 also shows correlations across states between the factor loadings for individual sectors and those for total output. In general these are reasonably (but not overwhelmingly) strongly correlated: i.e., the ranking of the impact of both factors across states is fairly similar across sectors. There is however one conspicuous exception: as discussed in Section 7 the weightings of V-factor loadings for the public sector are inversely correlated with those for output: ie, V-States typically had inverted V-factor patterns of public spending, and vice versa.

Note that the data for the period 1970-2000 do not pick out such a clear split between the states in terms of even the total V-Factor loading as do those for the longer sample 1960-2003. This is driven by the difference in sample, rather than the difference between total output and sectoral loadings.

Bibliography

Bai, Jushan (2003). "Computation and Analysis of Multiple Structural Change Models", Journal of Applied Econometrics, Vol. 18, 1-22.

Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data, *Journal of Econometrics*, 122, 137-183.

Bai, Jushan & Ng, Serena (2002). "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, vol. 70(1), pages 191-221

Bai, Jushan & Ng, Serena (2004). "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, vol. 72(4), pages 1127-1177

Bai, Jushan , and P. Perron, (1998). "Estimating and Testing Linear Models with Multiple Structural Changes," Econometrica, Vol. 66, 47-78.

Campbell, John (1994) "Inspecting the Mechanism", *Journal of Monetary Economics*

Chow, Gregory, C. (1960). "Tests of Equality Between Sets of Coefficients in Two Linear Regressions", Econometrica, 28(3), p. 591-605.

Barro, Robert and Xavier Sala-i-Martin (1992). "Convergence", *Journal of Political Economy*, 100(2), p. 223-251.

Balakrishnan, Pulapre, and M. Parameswaran (2007). "Understanding Economic Growth in India: A Prerequisite," *Economic and Political Weekly*, July 14 Edition, p. 2915-2922.

Clark, Gregory and Susanna Wolcott (2003). "One Polity, Many Countries: Economic Growth in India, 1873 - 2000," in In Search of Prosperity: Analytic Narratives on Economic Growth, ed. Dani Rodrik. Princeton, New Jersey: Princeton University Press.

Dasgupta, Dipankar, Maiti, Pradeep, Mukherjee, Robin, Sarkar, Subrata, and Subhendu Chakrabarti (2000). "Growth and Interstate Disparities in India", *Economic and Political Weekly*, July 1 Edition, p. 2413 - 2422.

Datt, Gaurav and Martin Ravallion (2002). "Is India's Economic Growth Leaving the Poor Behind?," *Journal of Economic Perspectives*, Vol 16(3), p. 89-108.

Ghate, Chetan (2008). "Understanding Divergence in India: A Political Economy Approach", *Journal of Economic Policy Reform*, In Press.

Hansen, Bruce, E. (2001). "The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity," Journal of Economic Perspectives, 15(4), p. 117-128.

Hausmann, R., Pritchett, L., and Dani Rodrik, (2005), Growth Accelerations, Journal of Economic Growth 10, p. 303-329.

Mankiw, Gregory Romer, David and David Weil (1992). "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, Vol. CVII, p. 407-37.

Nagaraj, Rayaprolu, Varoudakis, Aristomene, and Marie-Ange Vengazones (1998). "Long Run Growth Trends and Convergence Across Indian States," OECD Technical Papers, p. 1-58.

Rodrik, Dani and Arvind Subramanian. (2005). "From Hindu Growth to Productivity Surge: They Mystery of the Indian Growth Transition." IMF Staff Papers, Volume 52(2).

Trivedi, Kamakshya (2002). "Regional Convergence and Catch-up In India between 1960 and 1992," Mimeo, Oxford University.

Virmani, Arvind. (2006), "India's Economic Growth History: Fluctuations, Trends, Breakpoints, and Phases", *Indian Economic Review*, Vol. 41(1), p. 81-103.

Wallack, J. (2003), Structural Breaks in Indian Macroeconomic Data, Economic and Political Weekly, October 11.