States Diverge, Cities Converge:
Drivers of Local Growth Catch-Up in India

Yue Li, Martín Rama and Qinghua Zhao *

This draft: May 22, 2017

Abstract

There is broad consensus that growth within India is characterized by "divergence, big time". We take a fresh look at this proposition, departing from the existing literature in three ways. First, we assess growth patterns across districts and across places below the district level, instead of taking the state as the unit of analysis. Second, we rely on household expenditures per capita to measure living standards, instead of GDP per capita. And third we use a Bayesian model averaging approach to identify the key drivers of local growth, instead of using OLS. In contrast with the growth divergence consensus, we find that living standards strongly converge across districts and places below the district level, with growth being fastest in the mid-range between purely rural places and major urban centers. Divergence at the state level is most possibly due to the fact that low-income states do not generate enough of those fast-growing locations. Our findings on drivers of local growth also depart from conventional wisdom. Access to electricity, transport infrastructure and market access come in very strongly, but not irrigation or housing investments. The coverage of primary education is an important predictor of growth, but not that of other levels of education. The sectoral structure of economic activity does not seem to play a role, but locations with a bigger share of medium and large firms grow substantially faster. Social inclusion—access to finance, gender equality and social homogeneity—matters, and geography too. On governance, law-and-order at the state level makes a difference, as do state-level labor and land regulations, but variations in city governance come out as only mildly relevant. The prominence of state-level governance may explain why low-income states are failing to generate more vibrant locations.

---

* Yue Li and Martín Rama are with the World Bank's office of the Chief Economist for South Asia; Qinghua Zhao is with its Development Research Group. The findings, interpretations, and conclusions expressed in this paper are entirely theirs, and do not necessarily represent the views of the World Bank. The corresponding author is Yue Li, at yli7@worldbank.org. The authors gratefully acknowledge the skillful research assistance provided by Virgilio Galdo, María Florencia Pinto, and Jiarui Wang, and the useful comments and suggestions received from Urmila Chatterjee and Rinku Murgai.

1. Introduction

India's growth performance over the past three decades has been nothing short of remarkable. Between 1985 and 2015, GDP per capita grew on average at 4.5 percent per year. Growth even accelerated within this period, from 3.1 percent in 1985-94 to 4.4 percent in 1995-2004, and surged to 6.1 percent during 2005-15 (World Bank, 2017). At the national level, India has been quickly catching up with more advanced economies, defying the "divergence, big time" claim (Pritchett 1997). Growth patterns at the subnational level are less encouraging. Income disparities within India are wide, to the point of justifying a distinction between high- and low-income states. And a key question is whether poorer locations are catching up with the richer ones as the country experiences rapid growth.

The consensus among the economics profession is that within itself India is indeed characterized by "divergence, big time". Datt and Ravallion (2002) find that the states that were richer in the 1980s grew faster in the 1990s. Kumar and Subramanian (2012) document not only a continuation of state-level divergence in the 2000s, but an increase in its magnitude. In the most positive assessment, Ghate and Wright (2013) conclude that there is no clear evidence in favor of absolute convergence or absolute divergence across states from the 1980s to the 2000s. The most recent *Economic Survey* by the government of India features this convergence debate, showing the sharp contrast with growth convergence within China and and other large federations (MOF 2017).

However, the same *Economic Survey* also points to new evidence that non-monetary measures of wellbeing are improving across all states, and especially in low-income ones (MOF 2017). It also reports higher levels of cross-state transactions and cross-state migration than previously documented. In contrast with previous research, a recent study using the intensity of nightlights to proxy for income finds clear evidence of convergence across Indian states, districts and cities (Tewari and Godfrey 2016).

In this paper, we take a fresh look into the Indian convergence debate, departing from the existing literature in three important ways:

- We assess the growth patterns across districts and across places below the district level. Some Indian states are bigger than all but a few countries in the world. Taking the state as the spatial unit of analysis can hide substantial heterogeneity in local growth patterns.
- We rely on household expenditures per capita to measure living standards. This indicator, estimated out of household expenditure surveys, provides greater spatial granularity than GDP per capita, while still being highly correlated with it.
- We explicitly identify what makes some locations grow faster than others. One important challenge in doing so is the multiplicity of potential drivers of growth, most of which are correlated with each other. We address this challenge by using a Bayesian model averaging approach.

In contrast with the growth divergence consensus, and in line with more indirect evidence provided by non-monetary indicators of wellbeing and nightlights intensity, we find a strong convergence in living standards across districts and places below the district level between 2004 and 2011. On average, a district whose household expenditure per capita in 2004 is 10 percentage points lower experienced an annual growth rate that was about 0.25 percentage points higher. This result holds when considering all states, and also when restricting the analysis to the largest states. And the result is not driven by smaller districts and places with potentially larger measurement errors.

Importantly, there is also convergence in location premiums. These are the spatial differences in household expenditures per capita after controlling for the characteristics of the households living in those locations – their demographics, educational attainment and assets. Location premiums can be interpreted as a local productivity measure. Convergence in household expenditures per capita could be driven by migration involving distinct subsets of households. For instance, household expenditures per capita would increase in locations that attract more educated migrants. But convergence in location premiums implies that there is more than self-selection of households at play, and a catch-up in productivity is actually taking place.

The Bayesian model averaging approach allows us to screen the most important drivers of this convergence process. We start with 30 potentially important initial conditions, identified based on previous literature on convergence across countries and across cities. These initial conditions cover nine plausible "buckets" of determinants, from connectivity to skills and from governance to economic structure. However, several of the indicators are correlated with each other, and as a result their estimated impacts on local growth depend on which other indicators are being controlled for.

Given the large number of potential indicators, there are billions of possible regressions to be run. For instance, with 30 indicators plus the initial level of household expenditure per capita, up to 2.1 billion models can be considered (= 2^31). The Bayesian model averaging approach uses a rigorous methodology to select which indicators to include in the analysis. In practice it amounts to estimating these billions of regressions starting from some prior regarding which variables to include, and adjusting those priors along the way to improve the fit of the regressions.

Until recently, computational capabilities were a constraint to run this kind of analysis, forcing researchers to use approximations. One of the contributions of this paper was to develop an algorithm allowing to run one billion regressions on an ordinary laptop in less than one hour. The C++ code is available on request.

Previous attempts to identify drivers of growth through regression analyses have been met with skepticism. However, the three distinctive features of this paper mitigate the main criticisms:

- One reason for skepticism regarding this literature is the limited number of countries relative to the potential number of drivers of growth. The smaller the number of degrees of freedom left, the more unstable the estimated coefficients are (Levine and Renelt 1992). This criticism is all the more valid when considering convergence across states in India, as there are three dozen observations at most. However, by conducting the analysis at the district or below-district level, we can rely on many hundreds to more than one thousand observations, thus removing this concern. Degrees of freedom are not a constraint in our case.

- Another reason for skepticism is the sensitivity of results from the Bayesian model averaging approach to small changes in the data. It has been shown that revisions to the GDP data used in cross-country growth regressions dramatically alter the selection of variables considered relevant (Ciccone and Jarocinsky 2010). By using detailed survey data from hundreds of thousands of households, our income estimates are arguably more stable than sub-national GDP figures. And unlike national GDP figures, they do not suffer from the variability introduced by still-unreliable purchasing power parity adjustments (Deaton and Tten 2017).

- The Bayesian model averaging approach itself has been questioned because the probabilities of including specific drivers of growth in the analysis can be unstable (Ciccone and Jarocinsky 2010).

> Given that we are not constrained by degrees of freedom, it could be argued that a safer approach would be to include all the potential drivers of growth. However, while the probability to include a particular driver of growth varies depending on priors and sample, the average coefficients across billions of regressions are remarkably stable. And they are different from the coefficients obtained when running a single regression including all the potential drivers of growth.

In light of the above, we are reassured that our results are robust. Depending on the statistical criteria used, 12 to 21 of the initial conditions are retrained as robust predictors of subsequent economic growth at the local level.

While some of these results confirm the findings of previous research, others contradict conventional wisdom. Thus, access to electricity, transport infrastructure and market access come in very strongly, but not irrigation or housing investments. Similarly, the coverage of primary education is an important predictor of growth, but not that of tertiary education. On governance, law-and-order at the state level matters, as do state-level labor and land regulations, but variations in city governance come out as only mildly relevant. We do not find evidence that the sectoral structure of economic activity (say, the share of manufacturing) plays a role, but locations with a bigger share of large firms perform substantially better. Importantly, all indicators of inclusion considered – higher access to finance, lower gender gaps in educational attainment and greater social homogeneity – strongly contribute to local growth. And geography also matters, with both elevation and rainfall having a significant impact. These findings call for a reconsideration of commonly held priors about India's economic growth.

The findings also show, convincingly we believe, that there is growth divergence at the state level but growth convergence when districts, or places below district level, are considered. Growth is fastest in the mid-range between purely rural places and major urban centers. But the top performers do as well in high- and low-income states. It is just that there are not many top performers among the latter. The governance results in the analysis of the drivers of growth helps make sense of this finding. With state-level governance being more relevant than city governance, low-income states fail to create the conditions for the emergence of more vibrant cities.

2. Previous research

Studies on economic convergence differ in the way convergence itself is defined, in their unit of observation, and in their statistical approach. On the definition, an important distinction is usually made between absolute and conditional convergence. Absolute convergence analyses focus on whether administrative units with a low initial income per capita grow faster than those with a high initial income. Conditional convergence analyses also control for other factors that could affect the speed of growth, in addition to the initial level of income per capita.

Initially, the administrative unit considered in this literature was the nation state and the performance indicator was GDP per capita measured in comparable purchasing power terms (Baumol 1986, Barro 1991). This literature started with initial GDP per capita, the investment rate, and secondary school enrollment as the key drivers of growth. But the number of potentially relevant indicators included in the specification burgeoned over just a few years, reaching the point where degrees of freedom started to become scarce. Not surprisingly, there was substantial uncertainty as to which potential drivers of growth mattered most. But there was consensus that the world was experiencing both absolute divergence and conditional convergence.

Building on this strand of the literature, studies started focusing on convergence across "clubs" of similar countries, within supra-national federations, and across sub-national administrative units (Barro and Sala-i-Martin 1991, 1992, 2004). While specifications are similar to those used in the cross-country literature, a notable finding is the switch from absolute divergence to absolute convergence. This result holds for the 48 states of the United States during 1880-2000, the 47 prefectures of Japan during 1930-90, and the 90 regions of eight European countries during 1950-90. Subsequent studies confirmed this finding for Australasia, Canada, Ireland, and Sweden.

Meanwhile, the urban literature had been exploring the determinants of city growth focusing on one particular set of potential drivers at a time. Glaeser, Scheinkman and Shleifer (1995) broke with this tradition by adapting the cross-country regression model to the analysis of growth across US metropolitan areas in 1960-90. The performance indicators were the city population and the city wage rate, with schooling, unemployment and the employment share of manufacturing identified as robust drivers of urban growth. -Other studies involving multiple indicators were conducted by da Mata et al. (2007) for Brazilian cities in 1970-2000, and by Duranton (2016) for Colombian cities in 1993-2010. The results of these other studies suggest that geography, road connectivity, educational attainment and economic specialization are important drivers of subsequent city growth.

In India's case, there is ample consensus that regional imbalances are wide. A state like Goa has a GDP per capita comparable to that of Mexico, while GDP per capita in the state of Bihar is closer to that of Benin. Large sub-national gaps in living standards, including both monetary and non-monetary dimensions of wellbeing, are also documented in multiple studies. But fewer analyses have focused on gaps at district or below-district levels.

Most convergence studies for India actually use the state as their unit of observation. The literature on this topic up to 2007 was summarized by Kalra and Sodsriwiboon (2010). Adding more recent publications, there have been 19 studies on sub-national convergence in India (table 1). These studies use different methodologies, including cross-sectional regression, panel regression, and distributional analysis. Bajpai and Sachs (1996), Rao et al. (1999), and Kumar and Subramanian (2012) split the sample into various time periods, which leads to multiple results. Kochhar et al. (2006) also report multiple results because they use different estimation methods.

Out of the 26 results, 18 support absolute divergence, two are inconclusive, and six favor absolute convergence. The evidence in favor of absolute divergence is strongest at the state level, and especially so in the 1980s and 1990s. Datt and Ravallion (2002) report absolute divergence across all states, except for the richest two in the 1990s. For the same decade, Rodrik and Subramanian (2004) find that states are diverging at an annual rate of about 1.2 percent a year during the periods, a result that they see as a striking case of "Divergence, Big Time".

Findings are more mixed for the 2000s, although on balance the evidence suggests that absolute divergence remained the norm. Kumar and Subramanian (2012) even report an increase in the magnitude of divergence in the 2000s in comparison with the 1990s. MOF (2017) expands the same analysis to 2004 for income, and to 2011 for expenditure, and confirms that there is continuing divergence across states. When using income as the performance indicator, Das (2012) also finds strong divergence across states in 1980-2005. However, Ghate and Wright (2013) find no clear evidence to support either absolute convergence or absolute divergence across major states in 1987-2004. In addition to assessing the correlation between initial income and growth directly, they conduct millions of convergence regressions controlling for other factors potentially affecting growth, and find that the coefficient on initial income is symmetrically distributed around zero.

Table 1 A survey of the literature on growth convergence within India

Study	Convergence Absolute	Convergence Conditional	Dependent Variable	Spatial level (number)	Time period
Cashin and Sahay (1996)	Convergence	Convergence	Income	State (20)	1961-1991
Bajpai and Sachs (1996)	Convergence	Convergence	Income	State (19)	1961-1971
	Divergence	Divergence	Income	State (19)	1971-1995
Rao and Kalirajan (1999)	Divergence	Divergence	Income	State (14)	1961-1991
Nagaraj et al. (2000)	Divergence	Convergence	Income	State (17)	1971-1991
Aiyar (2001)	Divergence	Convergence	Income	State (19)	1971-1995
Sachs et al. (2002)	Divergence		Income	State (14)	1976-1995
Datt and Ravallion (2002)	Divergence		Income	State (12)	1992-2000
Bandyopadhyay (2003, 2012)	Divergence		Income	State (17)	1965-1997
Rodrik and Subramanian (2004)	Divergence		Income	State (20)	1980-1999
Baddeley et al. (2006)	Divergence		Income	State (15)	1970-1997
Kochhar et al. (2006)	Inconclusive		Income	State (14)	1961-2000
Purfield (2006)	Convergence	Convergence	Income	State (15)	1976-2005
Misra (2007)	Divergence	Convergence	Income	State (14)	1976-2001
Kalra and Sodsriwiboon (2010)	Divergence		Income	State (15)	1960-2003
Das (2012)	Divergence		Income	State (14)	1980-2005
	Divergence		Rural expenditure	State (14)	1980-2005
	Convergence		Urban expenditure	State (14)	1980-2005
Kumar and Subramanian (2012)	Divergence		Income	State (21)	1993-2009
	Divergence		Income	State (21)	2001-2009
Ghate and Wright (2013)	Inconclusive	Inconclusive	Income	State (15)	1987-2004
Nayyar (2014)	Divergence	Convergence	Expenditure	State (17)	1994-2012
Das et al. (2015)	Divergence	Convergence	Income	District (575)	2001-2008
Tewari and Godfrey (2016)	Convergence		Nightlight intensity	State (33)	1992-2013
	Convergence		Nightlight intensity	District (618)	1992-2013
	Convergence		Nightlight intensity	City (479)	1996-2011

Note: Income stands for either GDP per capita or Gross Value Added per capita. Expenditure stands for average household expenditure per capita. Nightlight intensity is measured per unit of surface.

Source: Authors, building on Kalra and Sodsriwiboon (2010).

With few studies considering higher levels of spatial disaggregation, results at the district level or below-district level appear to be inconclusive. Das et al. (2015) find that richer districts in 2001 grew slightly faster during period 2001-2008. When using expenditure as the performance indicator, Das (2012) also finds evidence of divergence across the rural areas of states, but he reports convergence across their urban areas. This finding is consistent with the evidence presented by Tewari and Godfrey (2017). Using nightlight intensity as the performance indicator, this paper concludes that there is absolute convergence across states, districts and cities.

Based on the convergence literature, not just for India but also across countries, “clubs” of countries and sub-national units, we identify 49 potential drivers of growth. This is in addition to the initial level of income, which is the key indicator in the convergence literature. In practice, some of these 40 potential drivers of growth can be measured by more than one indicator. For example, there is a presumption that urban governance matters, but Duranton (2015) captures it through three different measures: institutional capacity, fiscal performance and overall performance. To be succinct, we report the indicators as one as long as they refer to the same idea.

These 49 potential drivers of local growth can be classified into nine conceptual groups or buckets:

Infrastructure play a central role both in the neoclassical growth theory and in the new growth literature (Solow 1956, 1957, Barro 1991). In the Indian context, almost all studies consider measures of infrastructure or investment. In some cases the emphasis is on access to quality infrastructure services – electricity, transportation, telecommunication and irrigation. In others, public or private investment are emphasized. These indicators are generally found to matter: greater access to infrastructure, better quality of infrastructure services, and higher levels of investment are all positively associated with growth. But significance varies across subsectors.

Market access is an indicator favored by the new economic geography (Krugman and Venables 1995, Fujita et al. 1999). The ability to easily reach dynamic centers is viewed as a booster for agglomeration effects, as it supports both a stronger demand for final products and a more affordable and diversified supply for inputs. Market access critically depends on connectivity. From an empirical point of view it is often measured as a weighted distance to nearby economic activity. More sophisticated measures are derived from gravity models estimated using trade data (Redding and Venables 2004). Dummy variables for land-locked states are used as well. Empirically, market access seems to matter.

Another group of possible drivers is related to *economic structure*. The development literature emphasizes structural transformation, from an economy dominated by agriculture to one where manufacturing and services are predominating, as a source of faster productivity growth. In India, areas with a greater share of agriculture experience slower growth, while growth is faster in areas with a greater share of services. This finding seems consistent with the idea of structural transformation. However, greater shares of (registered) manufacturing are also found to grow more slowly, which has been attributed to the high regulatory burden faced by these industries in India (Ghate and Wright 2013). The urban economics literature, in turn, emphasizes the diversification of economic activity. For example, Duranton (2016) finds evidence of a relationship between specialization and city growth in Colombia.

The role of the *employment structure* has been less studied by the convergence literature. However, in a country undergoing a massive demographic transition it is worth exploring whether the nature of the jobs available at the local level influences subsequent growth. The urban economics literature holds that employment density underlies agglomeration economies and that labor pooling is an important channel through which some local economies become more productive. This suggests that employment indicators could be a good predictor of local growth. In the Indian context, none of the studies reviewed finds the employment structure to be related to growth. However, Glaeser, Scheinkman and Shleifer (1995) shows that the unemployment rate matters for city growth in the United States.

Many of the local administrative entities used as units of observations in the analysis encompass both urban and rural areas. This makes *urbanization* a potentially important determinant of local growth. The forces of agglomeration make urban areas more productive than rural areas (Rosenthal and Strange 2004, Duranton 2015). Indeed, most capital accumulation and knowledge spillovers take place in cities. Ghate et al. (2013) and Das et al. (2015) confirm that higher urbanization rates are related to faster growth across Indian states and districts. However, greater density can also give rise to diseconomies of scale. Duranton (2016) provides some evidence in support of this opposite argument.

Human capital is the ultimate driver of economic dynamism in the so-called endogenous growth literature (Romer 1986). Almost all convergence studies include human capital indicators such as school enrollment rates or average years of schooling among the population. Several of the studies find that human capital is an important factor in explaining growth in India. However, the significance of literacy rate as a measure of human capital is debatable. Two of the studies find that a higher literacy rate is positively related to subsequent growth, but in three others it is statistically insignificant. Beyond India, the urban literature tends to use years of schooling and tertiary educational attainment rates to measure human capital and finds strong evidence that it impacts local economic growth.

Social inclusion is a more elusive but not less important concept. There is consensus that discrimination against population groups, as well as barriers that prevent them to access markets and services, ought to be detrimental to growth. But which groups and which barriers should be the focus of the analysis is less clear. In the Indian context, it makes sense to pay attention to social cleavages along caste, tribe and gender lines. As for potential barriers, financial exclusion and lack of assets could prevent households from undertaking rewarding investments in human and physical capital. And insufficient access to education and jobs by women could deprive society of their full potential. This said, the studies reviewed find weak or little evidence that social inclusion contributes to local growth in India.

The relationships between *governance* and growth are addressed by the literatures on democracy and development, on social capital and on corruption. But multiple aspects of governance could be relevant, making this a broad bucket. India is a highly decentralized country. While some states are characterized by relatively strong governance, a state like Bihar was described as "lawless" until the turn of the century, with its murder rate much above the national average. States are also influential on concurrent policies that have strong implications on local growth, such as land law and labor law. For instance, since the seminal work of Besley and Burgess (2004) there has been much debate on whether states with more pro-business labor regulations do better than those with a pro-labor orientation.

Finally, *geography* has also been found to be important for local growth. Altitude and weather are among the most common indicators in the reviewed studies. However, the literature does not show evidence in favor of their significant impact on growth in India. The exception is being landlocked, but this is arguably a proxy for market access more than a geographical indicator.

The literature on convergence within India thus provides some clues regarding the contribution these buckets make to local growth (table 2). However, these results should not be interpreted literally, as they are seldom derived from specifications including all the potentially relevant variables.

Table 2 Drivers of growth in the literature of convergence within India

Buckets	Indicators	Results		
		Positive significant	**Negative significant**	**Small or inconclusive**
Infrastructure and investment	Access to electricity	Das et al. (2015)		
	Electricity losses		Purfield (2006)	Kalra and Sodsriwiboon (2010)
	Transport costs	Considered by da Mata et al. (2007) but not in the Indian literature		
	Access to road Road density			Purfield (2006) Das et al. (2015)
	Access to ports	Ghate and Wright (2013)		
	Telephone lines	Kalra and Sodsriwiboon (2010)		
	Irrigated land		Das et al. (2015)	
	Private investment	Rao et al. (1999) Aiyar (2001) Purfield (2006) Kalra and Sodsriwiboon (2010)		Baddeley et al. (2006)
	Public investment Development expenditure	Rao et al. (1999) Kalra and Sodsriwiboon (2010) Baddeley et al. (2006)		Ghate and Wright (2013)
	Government revenue Public debt	Considered by Glaeser et al. (1995) but not in the Indian literature		
Market access	Market access			Das et al. (2015)
	Distance to agglomerations		Das et al. (2015)	
	Landlocked		Ghate and Wright (2013)	

(Continued)

Table 2 Drivers of growth in the literature of convergence within India (continued)

Buckets	Indicators	Results		
		Positive significant	Negative significant	Small or inconclusive
Economic structure	Share of agriculture		Bajpai and Sachs (1996) Purfield (2006) Ghate and Wright (2013)	Baddeley et al. (2006)
	Share of manufacturing Share of industry		Purfield (2006) Ghate and Wright (2013)	
	Share of services	Kalra and Sodsriwiboon (2010)		
	Diversification	Considered by Duranton (2016) but not in the Indian literature		
	Specialization	Considered by Duranton (2016) but not in the Indian literature		
	Agriculture productivity	Baddeley et al. (2006)		
Employment structure	Employment-population ratio			Baddeley et al. (2006)
	Unemployment rate	Considered by Glaeser et al. (1995) but not in the Indian literature		
	Employment share of private sector			Kalra and Sodsriwiboon (2010)
Urbanization	Urbanization rate	Ghate and Wright (2013) Das et al. (2015)		
	City governance performance	Considered by Duranton (2016) but not in the Indian literature		
	City governance process	Considered by Duranton (2016) but not in the Indian literature		
Human capital	Literacy rate	Aiyar (2001) Ghate and Wright (2013)		Purfield (2006) Kalra and Sodsriwiboon (2010) Das et al. (2015)
	Tertiary educational attainment rate	Considered by Glaeser et al. (1995) and by Duranton (2016) but not in the Indian literature		
	Years of schooling	Considered by Glaeser et al. (1995) and by da Mata et al. (2007) but not in the Indian literature		

(Continued)

Table 2 Drivers of growth in the literature of convergence within India (continued)

Buckets	Indicators	Results		
		Positive significant	Negative significant	Small or inconclusive
Human capital	School enrollment rate	Baddeley et al. (2006)		
	Birth rate			Kalra and Sodsriwiboon (2010)
Social inclusion	Access to finance			Das et al. (2015)
	Female literacy rate			Purfield (2006)
	Gender gap in primary schooling	Baddeley et al. (2006)		
	Rural income inequality	Baddeley et al. (2006)		
	Urban income inequality			Baddeley et al. (2006)
	Income inequality Land inequality	Considered by Glaeser et al. (1995) but not in the Indian literature Considered by Duranton (2016) but not in the Indian literature		
	Segregation index	Considered by Glaeser et al. (1995) but not in the Indian literature		
Governance	Crime rate	Baddeley et al. (2006)	Rao et al. (1999)	
	Labor regulation index		Ghate and Wright (2013)	Purfield (2006)
Geography	Temperature	Considered by Duranton (2016) but not in the Indian literature		
	Precipitation			Ghate and Wright (2013)

Source: Authors.

3. Methodology

The basic specifications to assess convergence in the growth regression framework are:

$$g_{l,t,t+T} = \alpha^a + \beta^a y_{l,t} + \varepsilon_{l,t} \tag{1}$$

$$g_{l,t,t+T} = \alpha^c + \beta^c y_{l,t} + \Phi X_{l,t} + \mu_{l,t} \tag{2}$$

Equation (1) is used to assess absolute convergence and equation 2 for conditional convergence. $g_{l,t,t+T}$ is the annual growth rate of the living standards indicator in place l between t and $t+T$, and $y_{l,t}$ represents the level of the living standards indicator in l at t. $X_{l,t}$ is a vector of K other factors that potentially affect growth in place l. We use superscripts of a and c to differentiate the coefficients from the absolute and conditional convergence models. When $\beta^a < 0$, there is absolute convergence; when $\beta^a > 0$, there is absolute divergence. When $\beta^c < 0$, that is when other factors are controlled for, there is conditional convergence.

In this paper, the potential drivers of growth $X_{l,t}$ are assessed at or before t. This is admittedly not enough to ensure that they are fully exogenous. Except for indicators related to geography (altitude, precipitation and the like), the factors included in $X_{l,t}$ could be influenced by the anticipation of future growth in location l. However, the chosen time structure at least attenuates the risk of endogeneity.

An important methodological challenge faced by this framework has to do with the large number of variables that could potentially have an impact on local living standards. Both the theoretical and the empirical growth literature point to a multiplicity of factors that could make a difference, such as infrastructure, market access, human capital, governance and so on. In addition, there are several defensible indicators for each of these conceptual buckets. For instance, infrastructure may refer to access to electricity or to access to transport. As such, the number of potential factors K to consider is large. And the estimated impact $\emptyset^k$ of each of these indicators on the growth rate could vary depending on which other indicators are retained for the analysis. Given that the indicators are likely to be correlated with each other, a parsimonious approach could be picking up the effect of omitted variables.

One way to address this challenge is to run the analysis multiple times, including all the possible combinations of indicators, and to only retain as robust the indicators that remain significant in all circumstances, regardless of which other indicators are considered. However, this approach faces computational constraints, as the number of potential combinations increases exponentially with the number of potential drivers of growth. With K indicators, 2^K regressions would need to be considered. For example, our literature review identified 49 potential drivers of local growth, and that leads to 563 trillion possible combinations of indicators.

A practical alternative to make this approach tractable is to impose restrictions on the model, as Sala-i-Martin (1997) did. With his "two million regressions" method, he constrained equation (2) to a maximum of seven $X_{l,t}$ variables, with three of them fixed based on previous literature, and four flexible. This method reduces the number of possible regressions substantially. Even with 49 potential drivers of growth, the maximum number of regressions is "only" 163 thousand, the number of combinations of K – 3 taken by four. However, the method can be criticized on the arbitrariness of the restrictions imposed on equation (2).

An alternative with stronger analytical underpinnings is the Bayesian approach. The intuition in this case is that we do not know, among the many billion potential combinations of indicators, which one is the "true" model of the economy. We only have priors about it, but we accept to revise our priors based on the evidence. One appealing way to do this is the Bayesian Averaging of Classical Estimates, or BACE, proposed by Sala-i-Martin, Doppelhofer and Miller (2004). Their method applies "diffuse priors", under the form of probabilities that a particular indicator belongs in the true model.

Let M_j denote model j, defined as a subset of the K indicators that contains k_j of them, and let $P(M_j)$ be the prior probability attached to model M_j by the researcher. BACE runs equation (2) repeatedly, randomly drawing each time a subset of the K potential indicators based on prior probabilities $P(M_j)$.

After doing this a sufficiently large number of times, the posterior probability of individual model M_j is computed as:

$$P(M_j|Z) = \frac{P(M_j)N^{-k_j/2}SSE_j^{-N/2}}{\sum_{i=1}^{2^K} P(M_i)N^{-k_i/2}SSE_i^{-N/2}} \quad (3)$$

where Z represents the data, N represents the total number of observations, and SSE_j represents the sum of squared errors of model M_j. This is the equivalent of normalizing the weighted prior probability of each model by the sum of the weighted prior probabilities of all models, with weights determined by the goodness-of-fit of each model. This is similar to the Schwarz model selection criteria.

With this posterior model probability at hand, BACE derives the posterior inclusion probability of each indicator X^k as the sum of the posterior probabilities for all models that include this indicator:

$$P(X^k|Z) = \sum_{i=1}^{2^K} P(M_i|Z) * I(X^k \in M_i) \quad (4)$$

where $I(X^k \in M_i)$ equals one when X^k belongs to the set of indicators that define model M_i. And following Bayes rule, the posterior mean of the coefficient $\emptyset^k$ for each indicator X^k conditional on the inclusion of the indicator in the true model can be calculated as:

$$E(\emptyset^k|Z) = \frac{\sum_{i=1}^{2^K} \widehat{\emptyset_i^k} * P(M_i|Z) * I(X^k \in M_i)}{\sum_{i=1}^{2^K} P(M_i|Z) * I(X^k \in M_i)} \quad (5)$$

where $\widehat{\emptyset_i^k} = E(\emptyset^k|Z, M_i)$ is the OLS estimate for $\emptyset^k$ with the set of indicators that define model M_i. Essentially, this is a weighted average of all OLS estimates for $\emptyset^k$ where the weights are the ratios between the posterior probability of each model including X^k and the sum of the posterior probabilities of all models that include X^k.

The conditional posterior variance of the coefficient φ for each indicator (x) can be calculated as:

$$Var(\emptyset^k|Z) = \frac{\sum_{i=1}^{2^K} Var(\emptyset^k|Z, M_i) * P(M_i|Z) * I(X^k \in M_i) + \sum_{i=1}^{2^K} (\widehat{\emptyset_i^k} - E(\emptyset^k|Z))^2 * P(M_i|Z) * I(X^k \in M_i)}{\sum_{i=1}^{2^K} P(M_i|Z) * I(X^k \in M_i)} \quad (6)$$

This expression takes into account the variance of $\emptyset^k$ in each regression models, $Var\left(X^k \middle| Z, M_i\right)$, as well as the dispersion of the estimates for $\emptyset^k$ across all regression models, $(\widehat{\emptyset_i^k} - E\left(\emptyset^k \middle| Z\right))^2$.

The statistics of BACE are thus intuitively defined, and criteria similar to those of classical econometrics can be used for inference. Because of its relative simplicity and straightforward interpretation, BACE provides an effective approach to rank indicators based on their relevance, and to make judgements on the significance and stability of the coefficients attached to them.

In this paper we use prior model probabilities based on our review of the literature. As discussed in the previous section, we find that possible drivers of local growth fall into nine conceptual buckets and that for most of the buckets at least one indicator is found to be significant by previous studies. Therefore, a first way to define prior model probabilities is to classify the pool of potential drivers of local growth into the nine buckets and assume that one indicator from each group belongs in the true model. This way, all indicators from the same group have the same prior inclusion probability $1/k_b$ where k_b is the number of indicators in the bucket. It follows that indicators from different buckets have different prior inclusion probabilities, with their probability being lower the higher k_b is.

A second, more agnostic way is to assume that all indicators considered stand the same chance to belong in the true model, so that their prior inclusion probability is $1/K$. These two ways of thinking about prior inclusion probabilities are obviously very different. But as shown below, they both lead to the same set of posterior inclusion probabilities, which can be seen as proof of the robustness of the approach.

We also depart from the previous literature on the computational front. Sala-i-Martin, Doppelhofer and Miller (2004) run 100 thousand regressions at a time, at which point they replace the prior probabilities with the estimated posterior probabilities and start again, until the process converges and probabilities stabilize. A compact algorithm developed for this paper allows us to run one million regressions at a time, which helps the process converge on its own, minimizing the replacement of probabilities along the way. We test our computation process on the original dataset used by Sala-i-Martin, Doppelhofer and Miller (2004). Our results are highly consistent with their baseline results.

4. Data

The literature on convergence within India largely follows the growth regression framework, applying it mostly at the state level. Only a few studies go down to the district level, and only one applies the framework to cities. In this paper, by contrast, we consider three types of locations: states, districts and places. The latter represent a breakdown of the district into up to four types of agglomerations: small rural (less than 5,000 inhabitants), large rural, small urban (less than one million inhabitants) and large urban. This breakdown was first introduced by Chatterjee et al. (2015), and its use in a different context has shown its relevance to account for differences in living standards (Li and Rama 2015b).

Most of the studies for India use GDP per capita as their living standards indicator, with only a few considering household per capita expenditures, and only one focusing on nightlights intensity. In this paper we use household per capita expenditures as the common living standards indicator across all three types of locations. Data are from the Household Consumer Expenditure modules of the 61st and 68th rounds of National Sample Surveys of India, hereafter identified as NSS 2004-05 and NSS 2011-12. These surveys were conducted between July 2004 and June 2005, and between July 2011 and June 2012, respectively (NSSO 2005, 2012). They report household consumption information on an itemized form.

We use monthly consumption based on the mixed recall period and divide by household size to compute monthly nominal per capita household expenditure. To allow for comparability across years, we deflate the resulting nominal expenditures by the nationwide consumer price index (World Bank 2017).

For the district-level analysis, we calculate the average per capita consumption expenditures for each district in 2004 and 2011 and compute the annual growth rate. The procedure is the same for state-level figures. To generate performance measures at the place level we follow the approach developed by Chatterjee et al. (2015). This approach exploits the fact that NSS 2004-05 and 2011-12 follow a stratified multi-stage sampling design. Each district of a state or union territory is stratified into rural and urban areas. Within each stratum, first-stage units are ordered by their population and then further stratified into small rural, large rural, small urban and large urban places.

When using household expenditures per capita as the living standards indicator, instead of GDP per capita, the findings can be contaminated by migration. If migrants from low-income locations work in more vibrant locations and send remittances back home, they generate GDP in the latter but support consumption in the former. In India, migrants are more likely to be of working age, and more likely to have to have attended school, than the family members who stay behind (Kundu and Sarangi 2007, Rama and others 2014). The sending location may therefore be growing more slowly than the receiving location because its population has a shrinking share of relatively productive individuals, not because the location itself is becoming less productive in any fundamental way.

To account for this possibility, we consider the growth in location premiums as an alternative performance indicator. The location premium is the additional expenditure per capita a household enjoys because of the location it lives in. Put differently, the location premium is the variation in local expenditure per capita after controlling for household characteristics (Li and Rama 2015b). Because household characteristics are "removed", the annual growth rate of the location premium provides a measure of local productivity growth. We estimate location premiums at both the district level and the place level.

As regards the potential drivers of local growth, following our literature review we consider nine conceptual buckets and compute 30 indicators in all. In doing so we draw primarily from the *Spatial Database for South Asia*, compiled by the World Bank (Li et al. 2015). Of these indicators, 27 are district characteristics and three are state characteristics (the metadata can be found in appendix 1).

For *infrastructure*, we include four indicators, covering electricity, transportation, irrigation and housing. One caveat is that we use the density of railway and metro stations as our transportation measure. An alternative would be rely on road density, and an indicator based on Open Street Maps (a crowdsourcing platform) is available. However, the data are still too recent and unreliable for our analysis.

We construct two measures to approximate *market access* for each district. The first one uses GDP as an indicator of economic activity, the other one relies on nightlights intensity. For both measures we compute the average of all neighboring districts discounted by distance, for up to 400 kilometers. Unfortunately, we do not have bilateral district trade data to construct a theory-consistent measure following Redding and Venables (2004).

We collect seven indicators to capture *economic structure* of each district. Following the India convergence literature, we include the traditional measures on the sectoral composition of the economy. Following the urban economics literature, we introduce a measure on diversification. We also calculated the specialization index, but found that it was highly correlated with diversification index and therefore excluded it from the analysis. In addition, we introduce an indicator on mineral production capacity. The presence of mineral wealth is indeed believed to distort incentives, undermine competitiveness and dampen entrepreneurship, even in advanced economies (Glaeser et al. 2013).

While the indicators above refer to sectoral dimensions of the economic structure, recent literature on productivity finds that the structure in terms of types of firms matters as well (Melitz 2003, Hsieh and Klenow 2009, Bartelsman, Haltiwanger, and Scarpetta 2013). Even within narrowly defined sectors, productivity varies substantially across firms. When resources are allocated efficiently, productive firms grow quickly and less productive ones exit the market. Consequently, larger and younger firms tend to be more productive and tend to dominate the market, but this creative destruction process is often weak in developing countries. A considerable faction of firms remains small while many large firms are stagnant (Li and Rama 2015a). In the case of India, there is a long-standing debate about the missing middle in the distribution of firms, and its dampening impact on productivity growth (Tybout 2014). To capture the local strength of the creative destruction process, we introduce two measures of the distribution of firms by size, both based on the Economic Census.

Regarding the *employment structure*, we try to capture differences in the types of jobs available at the local level. In a developing country such as India, most people work, because they cannot afford to remain idle. Working part time on a farm or being self-employed is often one of the few options they have, even if they would prefer to have regular wage jobs. As a result, the unemployment rate tends to be low, and a large gray area between work, inactivity and unemployment exists. But the share of self-employment and wage employment relative to the working-age population, may provide a snapshot of the local employment structure.

On *urbanization*, it is important to keep in mind that the definition of urban varies widely across countries and is subject to debate in the Indian context. Using high resolution satellite imagery, some studies find the urbanization rate in India is underestimated by the administrative definition. To mitigate this concern, we use population density instead of the official urbanization rate. But because of this choice, we do not include in the analysis population size or population growth, two indicators which have at time been considered in the convergence literature.

Regarding city governance, Indian local urban bodies are generally considered weak in terms of their autonomy, resources and capacity. Consequently, their performance is poor on average and their influence over local economy is limited (MOF 2017). The exceptions are state capitals, who thanks to the federal Indian system enjoy substantial state autonomy and are empowered to coordinate decisions by various government departments. Therefore, we expect the presence of a state capital in a district to matter and use the population shares of these cities in a district to proxy for city governance.

For *human capital*, we exclude literacy rate because of its instability in previous studies. We also do not use school enrollment rates as it is not a measure of outcome. Instead, we include three measures on educational attainments, primary, secondary and tertiary. They are outcome variables and close to the concept of years of schooling that are often used in the city growth literature. Additionally, by differentiating between the three levels of education we allow nonlinearity in the impact of schooling.

To capture the multiple aspects of *social inclusion*, as the literature suggests, we include four substantially different indicators in the analysis. They capture the fraction of households with a bank account, the gap in educational attainment between men and women, and the shares of scheduled castes and scheduled tribes in the total population. Overall, our indicators thus cover issues such as financial inclusion, gender equality and social homogeneity.

On *governance*, we restrict our attention to state level characteristics. The main reason for this choice is the federal nature of the Indian institutional architecture, but has not so far led to a strong role by local bodies. States are responsible to enforce the rule of the law, but in addition they have a concurrent responsibility in relation to important laws, including those on labor and land. Following the literature, we use the average crime rate over a period of three years to capture law and order. We also use the labor regulation index introduced by Besley and Burgess (2004) and further developed by Aghion et al. (2008) to capture labor market rigidity. Similarly, we use the land reform index introduced by Besley and Burgess (2000) to assess the regulation of the land market.

For *geography*, we follow the literature and measure climate factors. To reduce the impact of short-term variations, we use decadal averages for temperature and precipitation that precedes the initial year. Additionally, elevation is a preexisting condition that affects climate and have implications on the costs of trade, transportation, and migration. Therefore, we also include it in our pool of indicators.

In our baseline analysis, we restrict our sample to the 21 largest states or union territories, referred to as “large states” in what follows (table 3). Most of the studies on convergence within India have focused on these large states, so our choice allows meaningful comparisons with previous findings. An additional reason for our choice that that two measures on state policies are only available for these 21 largest states. However, we do expand the exercise to 31 states or union territories to check the stability of the baseline results (appendix 2). All variables are standardized such that the coefficient estimates are comparable across regressions.

The correlations between some these indicators are high. The absolute value of 102 of their pairwise correlations is above 0.3, and 47 of the pairwise correlations are above 0.4 (appendix 3). Consistent with our expectations, 26 out of the 30 indicators are significant correlates to the growth of location premiums, after accounting for initial performance, based on the classical OLS estimates (appendix 4). This high correlation makes them indeed possible drivers of local growth.

5. Convergence, big time

Consistent with the literature, we find evidence in favor of absolute divergence of living standards across large states. This is made clear by splitting states into four groups, depending on whether they are above or below the median value of initial per capita expenditures and the median value of growth rates. It then appears that the majority of states fall into the right upper quadrant and the left lower quadrant (figure 1). A more rigorous OLS regression shows that the coefficient β^a in equation (1), which links initial per capita expenditure and the subsequent growth rate, is estimated at 1.5 and is statistically significant (table 4). This means that when the initial per capita expenditure of the state is ten percent higher, the state’s growth rate is about 0.15 percentage point higher.

Even if we expand the sample to 31 states, we find no evidence to support absolute convergence in living standards across states. The estimated coefficient β^a becomes negative but is statistically insignificant. The change in the sign of the relationship is primarily driven by the existence of three small rich states that grew relatively slowly.

Table 3 Summary statistics for the sample of large states

Performance	Indicators	Obs.	Mean	Std. dev.	Min	Max
Annual growth rate	Annual growth of per capita expenditure, district (percent)	512	11.48	3.02	2.52	24.40
	Annual growth of per capita expenditure, place (percent)	1218	11.26	3.88	-1.82	26.15
	Annual growth of location premium, district (percent)	512	9.83	2.54	3.07	18.41
	Annual growth of location premium, place (percent)	1218	9.96	2.54	1.43	18.41
Initial performance	Initial per capita expenditure, district (2004 India Rupees per month)	512	959	261	428	2205
	Initial per capita expenditure, place (2004 India Rupees per month)	1218	1017	316	419	3191
	Initial location premium, district (2004 India Rupees per month, log)	512	7.12	0.20	6.49	7.79
	Initial location premium, place (2004 India Rupees per month, log)	1218	7.16	0.18	6.49	7.79
Buckets	**Indicators**	**Obs.**	**Mean**	**Std. dev.**	**Min**	**Max**
Infrastructure	Access to electricity (share of total households)	512	0.52	0.28	0.03	0.98
	Railway station density (No. per 1000 sq.km.)	512	1.66	6.87	0.37	93.07
	Land with irrigation systems (share of total area)	512	0.28	0.26	0.00	0.96
	Share of housing in good condition (share)	512	0.47	0.15	0.15	0.91
Market access	Market access, nearby GDP weighted by distance (index)	512	4.22	2.35	0.56	13.63
	Market access, nearby nightlights weighted by distance (index)	512	0.45	0.04	0.37	0.55
Economic structure	Share of manufacturing (share of working age population)	512	0.06	0.05	0.00	0.38
	Share of other industries (share of working age population)	512	0.04	0.03	0.00	0.24
	Share of services (share of working age population)	512	0.13	0.06	0.02	0.37
	Share of medium size firms (share of firms)	512	0.0055	0.0033	0.0000	0.0250
	Share of large firms (share of total number of firms)	512	0.0008	0.0011	0.0000	0.0100
	Diversification index (index)	512	16	6	2.5	35
	Mineral production capacity (million metric tons per year)	512	58	482	0.4	5232

(Continued)

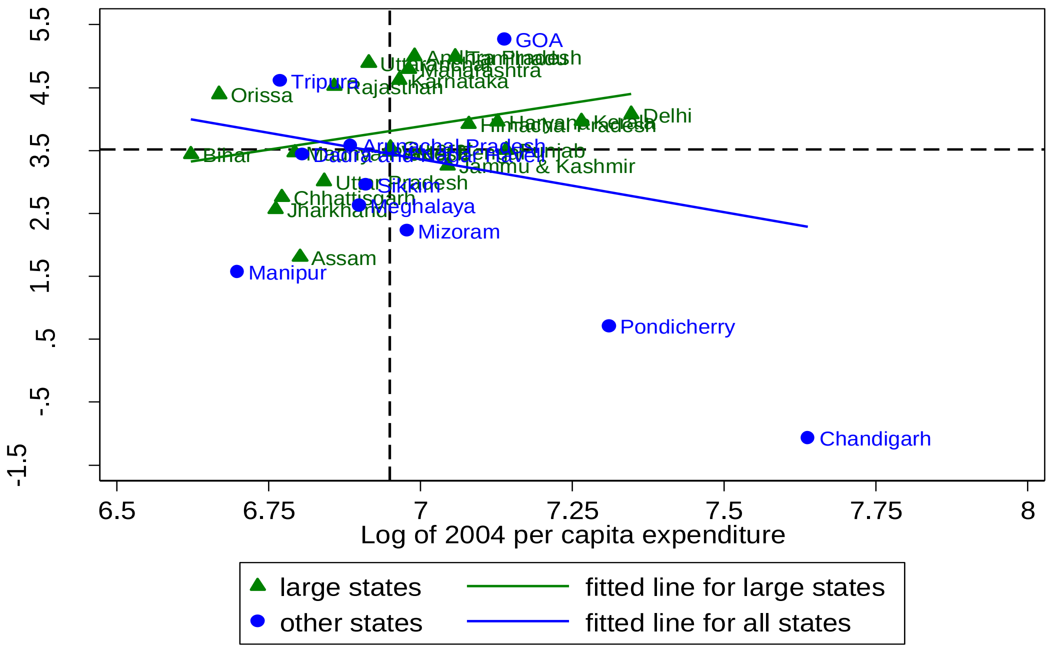
Table 3 Summary statistics for the sample of large states (continued)

Buckets	Indicators	Obs.	Mean	Std. dev.	Min	Max
Employment structure	Share of self-employed (share of working age population)	512	0.40	0.13	0.13	0.85
	Share of wage workers (share of working age population)	512	0.27	0.11	0.03	0.58
Urbanization	Population density (No. per sq.km.)	512	200	557	5	8995
	Population share of state capital (share of total population)	512	0.02	0.09	0.00	1.00
Human capital	Primary education (share of working age population)	512	0.17	0.07	0.02	0.43
	Secondary education (share of working age population)	512	0.04	0.03	0.00	0.19
	Tertiary education (share of working age population)	512	0.01	0.01	0.00	0.15
Social inclusion	Access to finance (share of total households)	512	0.33	0.14	0.07	0.80
	Gender gap in secondary education (percentage point difference)	512	0.03	0.02	-0.04	0.13
	Share of scheduled castes (share of total households)	512	0.13	0.23	0.00	1.00
	Share of scheduled tribes (share of total households)	512	0.19	0.11	0.00	0.54
Governance	Crime rate (No. of incidence per 1,000,000 people)	512	47	34	8	148
	Labor regulation index (index)	512	0.03	1.20	-2.00	4.00
	Land reform index (index)	512	4.51	3.03	0.00	15.00
Geography	Temperature (Degrees Celsius)	512	10	1	2	11
	Precipitation (mm.)	512	36	19	6	140
	Elevation (m.)	512	125	158	2	1339

Note: The 21 largest states or union territories are Andhra Pradesh, Assam, Bihar, Chhattisgarh, Delhi, Gujarat, Haryana, Himachal Pradesh, Jammu & Kashmir, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, Uttaranchal and West Bengal.

Source: Authors, based on Li et al. (2015).

Figure 1 Initial per capita expenditure and growth across states, 2004-2011



State level, all 31 states, median values defined by all states

Note: The green triangles indicate the 21 largest states and the blue dots the additional 10 states. The black dash lines indicate median values across the 31 states. The green line is the fitted linear regression for the 21 largest states and the blue line is for all 31 states.

In sharp contrast, we find strong evidence of absolute convergence of living standards when the district, rather than the state, is the unit of observation (table 4). The results are not sensitive to the choice of the sample, which is also in contrast to the results of state level analysis. For both large states and all states, the estimated correlation coefficient between initial per capita expenditure and subsequent growth rate, β^a, is significant and negative (between -2.6 and -2.5). On average, when district per capita expenditure in 2004 is 10 percent higher, the subsequent annual growth rate is 0.25 to 0.26 percentage point lower.

The speed of absolute convergence is even twice as fast when considering the place as the unit of observation (table 4). Again, the results are not sensitive to our choice of sample. The estimate of β^a remains significant and negative, with its absolute value almost doubling. When per capita expenditure at the place level in 2004 is 10 percent higher, the subsequent annual growth rate of the place is 0.47 to 0.49 percentage points lower.

This finding is not a statistical artifact, driven by higher measurement errors when considering smaller locations. It is true that the number of households surveyed becomes smaller when the level of spatial disaggregation goes from state to district and further to place, and as a result the standard errors of the estimated per capita expenditure become greater. This type of measurement errors could indeed bias our estimates in the direction of greater apparent convergence. For instance, if expenditures per capita in a place are overestimated in the initial year, this base effect is likely to lead to an underestimation of the annual growth rate in subsequent years, and this could be misconstrued as convergence.

To mitigate the impact of measurement error, we divide both districts and places into 25 quantiles, based on their initial expenditure per capita. We then estimate the mean initial per capita expenditure and the mean annual growth rate for each quantile. These means should not suffer from serious measurement error because there are many districts and places in each quantile. We finally re-run equation (1) on these 25 observations, and the estimated β^{α} coefficient barely changes as a result (table 4, figure 2).

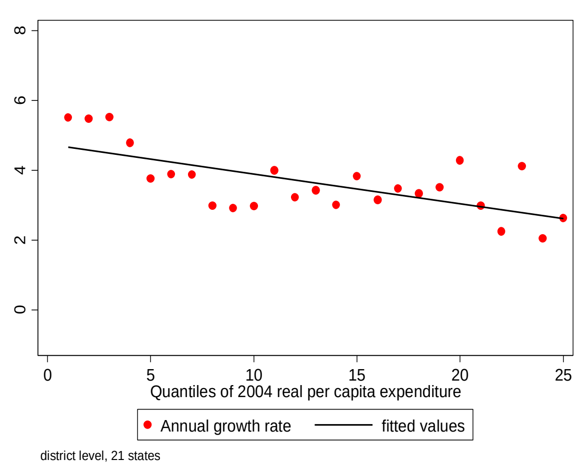
Table 4 Absolute convergence in per capita expenditure, 2004-2011

	State level		**District level**			**Place level**		
Sample	Largest 21 states	All 31 states	Largest 21 states	All 31 states	25 quantiles	Largest 21 states	All 31 states	25 quantiles
Log of 2004 per capita expenditure	1.479* (0.798)	-1.686 (1.866)	-2.574*** (0.540)	-2.537*** (0.513)	-2.472*** (-0.661)	-4.886*** (0.386)	-4.747*** (0.373)	-4.740*** (-0.503)
Constant	-6.469 (5.625)	15.161 (12.893)	21.233*** (3.685)	20.915*** (3.507)	20.468*** (-4.528)	37.203*** (2.644)	36.195*** (2.558)	36.149*** (-3.453)
No. of observations	21	31	512	561	25	1218	1315	25
R2	0.102	0.074	0.054	0.051	0.421	0.136	0.128	0.848
Adjusted R2	0.055	0.042	0.052	0.049	0.396	0.136	0.127	0.842

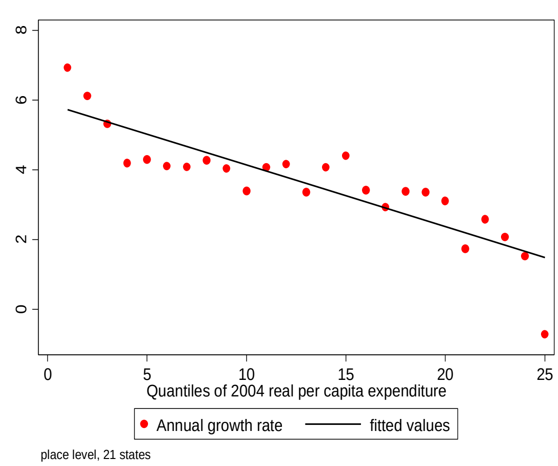
Note: Standard errors are in parentheses. Significance levels are: * 0.1, *** 0.05, and *** 0.01.

Figure 2 Initial per capita expenditure and growth across districts and places, by quantiles, 2004-2011

District level



Place level



The results hold when considering location premiums as the performance indicator, instead of household expenditures per capita (table 5). The estimated β^a coefficient at place level is comparable to that obtained when using expenditures per capita. But interestingly when focusing on location premiums, districts converge as fast as places. These results imply migration and the sorting of households across spatial units is not the driver of the observed convergence pattern. Other forces are at play and result in productivity growing faster in more disadvantaged locations in India during the second half of the 2000s.

Table 5 Absolute convergence in location premiums, 2004-2011

	District level			**Place level**		
Sample	Largest 21 states	All 31 states	25 Quantiles	Largest 21 states	All 31 states	25 Quantiles
2004 location premium	-4.170*** (0.561)	-4.362*** (0.534)	-4.130*** (0.495)	-4.698*** (0.386)	-4.804*** (0.366)	-4.876*** (-0.361)
Constant	39.603*** (3.996)	40.935*** (3.810)	39.321*** (3.505)	43.583*** (2.760)	44.340*** (2.620)	44.859*** (-2.581)
No. of Observations	512	561	25	1218	1315	25
R2	0.1112	0.1169	0.7121	0.1159	0.1189	0.8773
Adjusted R2	0.1094	0.1154	0.6995	0.1151	0.1182	0.8719

Note: Standard errors are in parentheses. Significance levels are: * 0.1, *** 0.05, and *** 0.01.

6. Why do states diverge?

The finding that districts and places converge, and quite rapidly, is of course much more encouraging than the prevailing "divergence, big time" consensus. But it also begs the question: if locations at lower levels of spatial disaggregation converge, why is that states do not? A defensible answer has to do with the distribution of fast-growing locations across states. One way to see this is to compute the share of fast-growing districts in each state, with fast-growing defined as having an annual growth rate above the national median. States comprising only one district are excluded from the analysis. In then appears that all states, including low-income states, have at least one fast-growing district. But low-income states have very few of these strong performers.

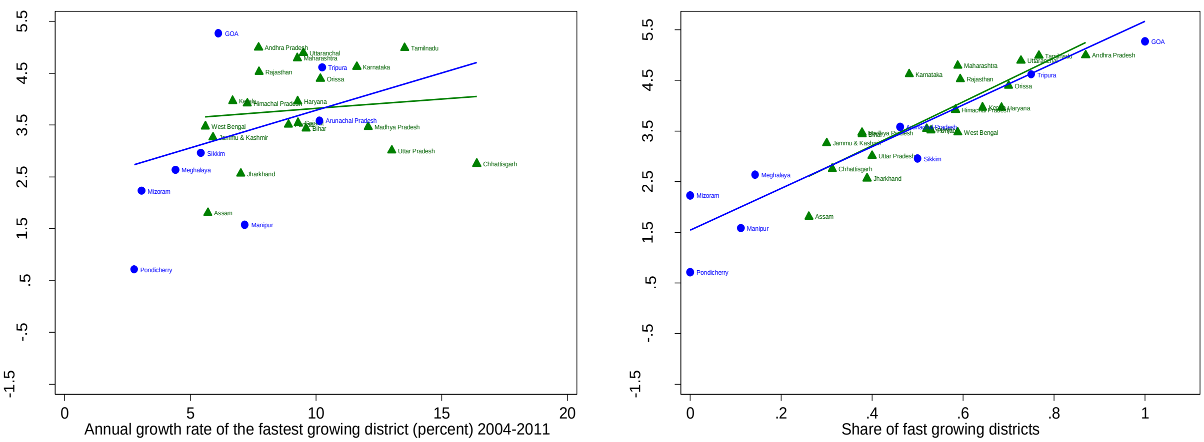
To illustrate the point, consider the annual growth rate of the fastest-growing district in each state. There is no clear correlation between the growth rate of the state and the growth rate of its fastest-growing district (figure 3). On average, all states have star districts that grow at par with peers in other states. But there is a positive correlation between the growth rate of the states and their shares of fast-growing districts. States with a higher share of fast-growing districts also tend to grow faster overall. Low-income states perform poorly because they do not generate enough of the fast-growing districts.

A simple regression analysis corroborates this result: the state-level growth rate is significantly and positively correlated with the share of fast-growing districts, but it is not correlated with the speed of the fastest-growing district (table 6). The results are stable to changes in the definition of fast-growing districts, such as considering the 75th percentile in the distribution of growth rates across districts as the relevant threshold. The results are robust to the inclusion of a measure of the share of rich districts in a state, with rich districts defined as those whose initial expenditure per capita is above the median value across all districts. The results remain valid if we expand the sample from large states to all states.

Figure 3 Fast-growing districts and growth of per capita expenditure across states, 2004-2011

Growth rate of fastest-growing district

Share of fast-growing districts



Note: Three states with only one district are excluded from the exercise. The green triangles represent the 20 largest states and blue dots indicate the additional eight states. The green line is the fitted linear regression for the largest states and the blue line is for all available states.

Table 6 Fast growing districts and growth of per capita expenditure across states, 2004-2011

	Largest states			**All states**		
Share of fast-growing districts (above median value)	4.397*** (0.649)		4.443*** (0.751)	3.930*** (0.402)		3.825*** (0.425)
Share of fast-growing districts (above 75th percentile)		5.364*** (1.292)			5.865*** (0.984)	
Growth rate of fastest-growing district (maximum)	0.054 (0.038)	-0.002 (0.053)	0.052 (0.042)	0.048 (0.032)	0.034 (0.046)	0.061 (0.035)
Share of rich districts (above median value)			-0.064 (0.476)			0.283 (0.345)
Constant	0.944* (0.526)	2.455*** (0.561)	0.973 (0.583)	1.235*** (0.291)	1.855*** (0.387)	1.013** (0.399)
No. of Observations	20	20	20	28	28	28
R2	0.733	0.510	0.734	0.826	0.654	0.831
Adjusted R2	0.702	0.453	0.684	0.812	0.627	0.810

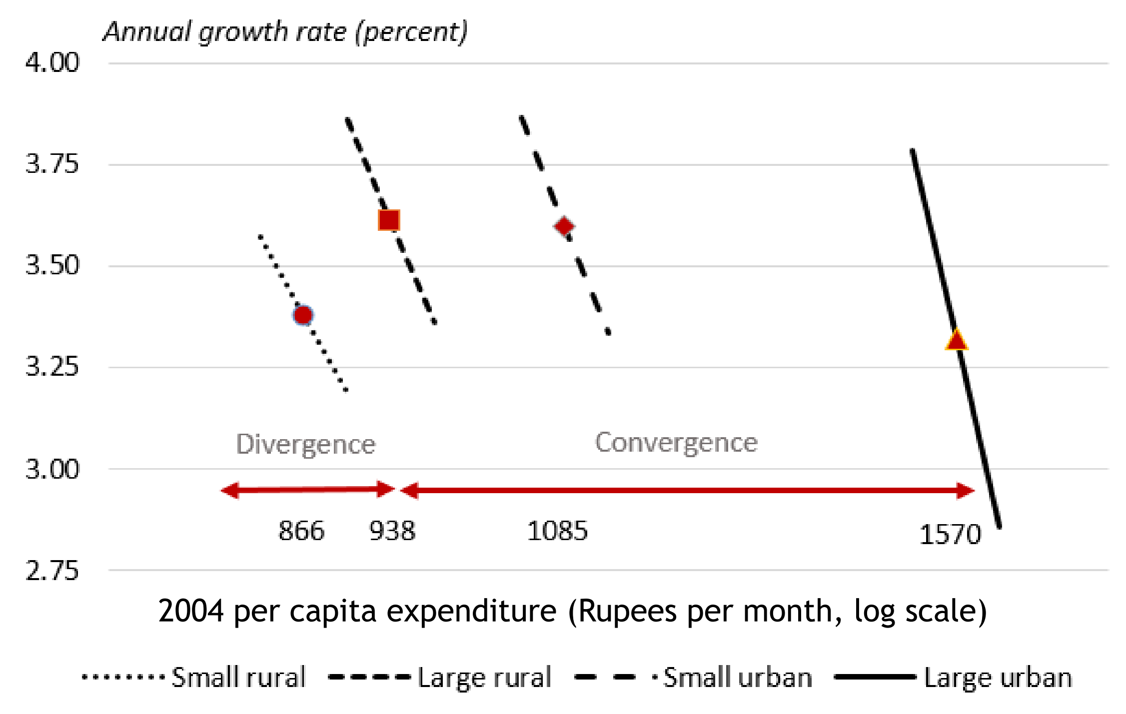
Note: Standard errors are in parentheses. Significance levels are: * 0.1, *** 0.05, and *** 0.01.

The conclusion is the same when we conduct the analysis at the place level instead of the district level, and when we use the location premium as the performance indicator instead of per capita expenditure (results are available on request). The magnitude of the positive association between the state growth rate and the share of fast-growing locations is largest when conducting the analysis at the place level with expenditure per capita as the performance indicator. It is smallest when looking at the district level and focusing on the location premium. But in no case do we find a statistically significant association between the growth rate of the state and that of its fastest growing location.

In sum, all states have fast-growing locations and even low-income states host locations growing at outstanding speed and catching up. However, low-income states are failing to converge because they face a shortage of these fast-growing locations. To put it differently the distribution of fast-growing locations is skewed toward rich states.

To understand what underlies this skewed distribution of local performance we explore the growth patterns of different types of places. We do this by re-running equation (1) separately for small rural, large rural, small urban and large urban places. We find that the estimated β^a is significantly negative across all four types of places, indicating strong convergence within each group. If anything convergence is strongest among large urban places. But average growth rates are higher in large rural and small urban places than in either small rural or large urban places (figure 4). The fast-growing locations identified above most often belong to this mid-range of the rural-urban gradation. All this suggests that the economic forces that sustain convergence are driven by the urbanization process. Low-income states may thus be failing to converge because they have not been as successful at urbanizing as other states.

Figure 4 Growth and convergence across the rural-urban gradation



7. Drivers of local growth

To identify the true model of the local economy, schematically represented by equation (2), we rely on the BACE approach. We use the growth of the location premium as the performance indicator and restrict the sample to the 21 largest states. The 30 indicators assembled on the basis of our literature review provide the set of potential drivers of growth. As discussed above, these indicators belong into nine conceptual buckets: infrastructure, market access, economic structure, employment structure, urbanization, human capital, social inclusion, governance and geography. Our diffuse prior is that the true model includes the initial performance of the location and one indicator from each of these nine buckets. Based on this diffuse prior, we draw one million random combinations of indicators from the nine buckets, and once this is done we replace our diffuse prior by the posterior probabilities estimated using BACE. We repeat the process until both the probability of including each indicator and its posterior conditional mean coefficient converge, which happens after seven rounds, or 7 million regressions.

The results are presented in table 7. Column (1) reports the posterior mean for each coefficient (β^c or $\emptyset^k$), conditional on the inclusion of the indicator in the model. This posterior mean coefficient, computed following equation (5), is a weighted average of all estimates of the coefficient from all of the regressions that include the indicator under consideration. The posterior model weights are computed based on equation (3). These weights reflect the goodness-of-fit of each model. These conditional posterior mean coefficients are comparable to coefficient estimates in standard regression analysis.

Because all indicators are standardized by subtracting their means and dividing by their standard deviations, the posterior mean coefficients in column (1) are comparable across indicators. The absolute values of a posterior mean coefficient can be interpreted as the relative impact of the indicator on growth. The ratio between the largest and the 10^{th} largest value of the posterior mean coefficients (both in absolute value) is about four, and the ratio between the largest and the smallest ones (also in absolute value) is about 430. These wide gaps suggest that the BACE approach discriminates effectively between indicators with very diverse marginal contributions to the goodness-of-fit of regression models.

The posterior inclusion probability for each indicator, computed following equation (4), is reported in column (2). The posterior inclusion probability is the sum of the posterior probabilities for all the regression models that include the indicator. It captures the weighted average of the goodness-of-fit of all those models. Not surprisingly, indicators with high posterior mean coefficients in absolute value also have high posterior inclusion probabilities.

The difference between the posterior and the prior inclusion probabilities is shown in column (3). We can divide the indicators according to whether their inclusion probability increases or decreases after running the BACE approach. For those indicators with increased or unchanged inclusion probabilities, our belief that they belong in the model is reinforced by the analysis. This metric provides a *first criterion* to determine which indicators to retain as part of the true model. Nine of the 30 indicators, regrouped at the top of table 7, meet this first criterion.

The posterior standard deviation for each coefficient conditional on the inclusion of the indicator, computed following equation (6), can be found in column (4). The conditional posterior standard deviation takes into account of both the variance of the estimated coefficient in each regression model as well as the dispersion of the estimates across regression models that include the indicator. The conditional posterior standard deviation is comparable to the estimated standard deviation in classical econometrics.

Table 7 Baseline results from Bayesian Averaging of Classical Estimates

Criterion	Rank	Variable	Posterior conditional mean coefficient	Posterior conditional inclusion probability	Posterior probability minus prior probability	Posterior conditional standard deviation	Posterior mean coefficient/ posterior std. deviation	Posterior t-statistic conditional on inclusion	Certainty of coefficient sign, in probability
			1	2	3	4	5	6	7
First	1	Initial location premium	-0.489	1.000	0.000	0.048	-10.193	-10.203	1.000
	2	Access to electricity	0.257	0.976	0.726	0.065	3.982	4.126	1.000
	3	Market access, lights	0.211	0.899	0.399	0.061	3.482	3.561	1.000
	4	Share of medium size firms	0.140	0.496	0.353	0.058	2.432	2.451	1.000
	5	Share of scheduled tribes	-0.131	0.488	0.238	0.053	-2.464	-2.494	1.000
	6	Elevation	-0.129	0.447	0.114	0.054	-2.398	-2.421	1.000
	7	Crime rate	-0.114	0.369	0.036	0.052	-2.175	-2.160	1.000
	8	Access to finance	0.113	0.264	0.014	0.056	2.024	2.028	1.000
	9	Labor regulation index	-0.110	0.457	0.124	0.047	-2.312	-2.308	1.000
Second	10	Primary education	0.108	0.226	-0.107	0.056	1.935	1.957	1.000
	11	Railway station density	0.104	0.222	-0.028	0.052	2.021	2.089	1.000
	12	Market access, GDP	0.100	0.111	-0.389	0.057	1.763	1.968	0.965
Third	13	Land reform index	0.105	0.144	-0.189	0.065	1.633	1.607	0.973
	14	Precipitation	-0.089	0.117	-0.216	0.056	-1.590	-1.651	0.998
	15	Tertiary education	-0.072	0.143	-0.190	0.049	-1.484	-1.482	1.000
	16	Diversification index	0.069	0.098	-0.045	0.053	1.287	1.292	1.000
	17	Share of large firms	0.066	0.091	-0.052	0.053	1.248	1.336	0.955
	18	Gender gap in secondary education	-0.054	0.083	-0.167	0.046	-1.171	-1.169	1.000
	19	Population share of state capital	0.049	0.068	-0.433	0.048	1.019	1.023	0.999
	20	Share of scheduled castes	-0.043	0.055	-0.195	0.051	-0.851	-0.843	0.978
	21	Mineral production capacity	-0.021	0.038	-0.105	0.045	-0.472	-0.471	0.990

(Continued)

Table 7 Baseline results from Bayesian Averaging of Classical Estimates (continued)

Criterion	Rank	Variable	Posterior conditional mean coefficient	Posterior conditional inclusion probability	Posterior probability minus prior probability	Posterior conditional standard deviation	Posterior mean coefficient/ posterior std. deviation	Posterior t-statistic conditional on inclusion	Certainty of coefficient sign, in probability
			1	2	3	4	5	6	7
Not retained	22	Share of self-employed	0.040	0.054	-0.446	0.055	0.737	0.721	0.932
	23	Share of housing in good condition	0.030	0.071	-0.179	0.066	0.449	0.393	0.625
	24	Secondary education	0.024	0.044	-0.290	0.059	0.414	0.427	0.701
	25	Share of wage workers	0.018	0.036	-0.464	0.059	0.306	0.298	0.773
	26	Population density	0.018	0.045	-0.455	0.060	0.296	0.383	0.611
	27	Share of other industries	0.013	0.035	-0.108	0.046	0.291	0.291	0.894
	28	Share of manufacturing	0.003	0.034	-0.109	0.049	0.071	0.074	0.578
	29	Land with irrigation systems	0.002	0.037	-0.214	0.059	0.030	0.038	0.456
	30	Temperature	0.002	0.055	-0.278	0.055	0.028	0.113	0.502
	31	Share of services	-0.001	0.037	-0.106	0.059	-0.019	-0.011	0.569

In the spirit of classical econometrics, we also calculate the ratio between the posterior mean coefficient and the posterior standard deviation. This ratio is akin to a t statistic. However, the conditional posterior t statistic can also be computed as the weighted average of t statistics across regression models, with weights defined as in equation (5). Columns (5) and (6) report these two ways of measuring the t statistic for each indicator. The two measures are highly consistent and provide a *second criterion* to define the significance of indicators. In classical terms, a coefficient would be 5-percent significant in a one-sided test if the absolute value of the t-statistic is greater than 1.66 for a sample size of 100 or above. Applying the 1.66 cutoff to both measures on t-statistics, we identify 12 indicators as significant. Not surprisingly, the nine indicators selected based on the first criterion are part of this set. The 12 selected indicators appear at the top of table 7.

Not all the estimates of a specific coefficient have the same sign as the mean posterior coefficient. Column (7) reports how likely the signs coincide, based on the posterior density. For each individual regression the posterior density of the coefficient is the same as in the classical regression model. In that model, a coefficient is significant at the 5 percent level if the probability of the true coefficient having the same sign as the estimated coefficient is 95 percent. This simple analogy provides us with a *third criterion* to identify the indicators that belong in the true model. Based on this criterion, 21 indicators are retained as statistically significant. They include the 12 indicators that are significant based on the previous two criteria. These 21 indicators are in the top portion of table 7.

Even by this least restrictive third criterion, the BACE approach allows us to discard at least one third of the indicators that economic theory or previous studies would have picked up as top candidates to drive growth at the local level. Additionally, if we rely on the two more strict criteria, we will discard almost half of the indicators. The method enables us to be selective about what matters.

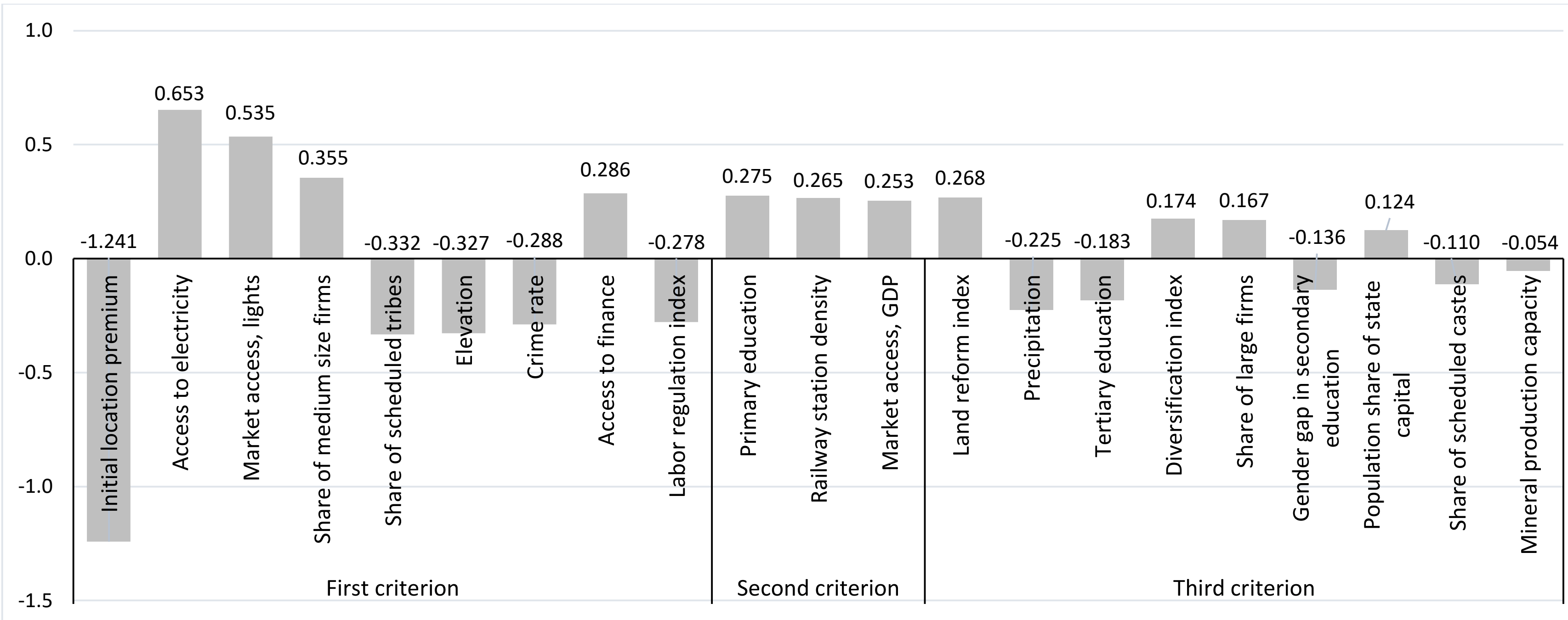
To facilitate the interpretation of the baseline results, we multiply the posterior means of the coefficients for the 21 significant indicators by the standard deviation of the growth rate of the location premiums. This multiplication yields the estimated difference in growth rates between two locations that would be identical in all respects, except that they would differ by one standard deviation in the value of the indicator of interest (figure 5).

Based on this transformation, the most relevant difference between two locations is the initial level of their location premium, an indicator that is significant according to all three criteria. This result reinforces the conclusion that India is experiencing convergence, big time.

The second most important set of indicators is related to infrastructure, and includes electricity and transportation. Access to electricity is identified as significant according to all three criteria and density of railway stations is identified as significant by two. The importance of infrastructure is consistent with the findings of the literature on convergence within India. However, our results also allow to differentiate between infrastructure services. In sharp contrast to the results on electricity and transportation, we do not find evidence that investments in housing and irrigation have a significant impact on local growth. This calls for a differentiated approach to infrastructure investments.

Market access also appears to be an important predictor of growth. Market access is identified as a significant indicator by all three criteria when it is computed based on nightlights, and by two criteria when it is computed based on GDP. This result is consistent with the findings of Redding and Venables (2004) and da Mata and others (2007). In the Indian context, it corroborates the claims by Ghate et al. (2015) that distance matters, and Li and Rama (2015b) that “distance to what” matters especially.

Figure 5 Robust predictors of growth



Note: The height of each bar represents the change in the annual growth rate, measured in percentage points, associated with an increase in the value of the corresponding indicator by one standard deviation.

While economic structure is found to have an impact on subsequent economic growth, it is not the sectoral structure that makes a difference but rather the size distribution of firms. In sharp contrast with previous studies, our results do not support retaining the shares of manufacturing, other industries or services as robust predictors of subsequent growth. Instead, we find locations with a larger faction of medium-size and large firms grow faster, as do places with a more diversified economic structure. The indicator on the share of medium-size firms meets all three criteria. The share of large firms, the diversification index and the mineral production capacity are all significant by the sign certainty probability criteria. Their impacts are more modest. These findings cast doubts on idea that specific sectors drive growth, much the same as previous research had cast doubts on the idea that specific sectors drive poverty reduction (Datt et al. 2016). Instead, these findings suggest that more attention should be given to firm dynamics and entrepreneurship.

Inclusion seems to be a solid predictor of faster growth. A one-standard deviation increase in the share of schedule tribes is associated with about 0.33 percentage point decrease in subsequent growth, a clear indication that this population group is being left behind. A one-standard deviation increase in the share of households with access to finance is associated with about 0.29 percentage points of additional growth. The gender gap in secondary education and the share of scheduled castes are also found to be significant by the sign certain probability criterion but their estimated impacts are much smaller. The finding that social inclusion is good for growth should not come as a surprise, but it marks a departure with previous studies on convergence within India, where inclusion is mostly absent.

Geography related factors also appear to be important for growth. Elevation is found to be negatively associated with growth and identified to be significant by all three criteria. The measure on precipitation is found to adversely affect growth and identified to be significant by one criterion. However, our results suggest that temperature does not have significant impact. This could be due to the fact that the specification in equation (2) is linear, whereas the literature on climate suggests the relationship between temperature and growth can be nonlinear.

Governance indicators measured at the state level are robust predictors of subsequent growth at the local level. The crime rate and the labor regulation index are significant indicators according to all three criteria. The land reform index is significant according to the sign certainty probability criterion. The crime rate, which can be interpreted as a measure of law and order, has a large impact. The impact is similar for the labor rigidity index, but much smaller for the land reform indicator. Overall, these results confirm the role of federalism in India, emphasizing the importance of strong governance at the state level.

Human capital also matters for growth, but the impact is not uniform across all education levels. Primary education completion is found to be significantly related to subsequent growth by two of the criteria. However, secondary education is found to be insignificant and tertiary education is found to be insignificantly by two of the criteria. While much emphasis is put these days in skills development and vocational training for adults, these results suggest that the early years are particularly important.

Regarding urbanization, we find weak evidence to support the influence of city governance and little evidence for population density. The measure on the population share of state capital is identified to be positive and significant by the sign certainty probability criterion. However, the magnitude of its impact is small. One possible interpretation of this result is city governance matters but the limited autonomy, resources and capacity of local urban bodies hinder their ability to exerting more influence on local growth. District population density is found to be insignificant by all three criteria. This result contradicts the conventional wisdom in urban economics that higher density equals stronger agglomeration effects.

Last but not least, we find little evidence to support the importance of employment structure. Wage employment and self-employment are identified as insignificantly related to subsequent growth by all criteria. This lack of evidence could be attributed to the fact that types of jobs in a location are intermediate outcomes, shaped by factors such as infrastructure, market access, economic structure and governance. Once these factors are adequately taken into account, the impact of the types of jobs available becomes negligible.

8. Robustness

The results above could be affected by the computational solution adopted to implement the BACE approach. One concern in this respect is whether the number of instances in which prior probabilities are replaced by posterior probabilities makes a difference. Sala-i-Martin et al. (2004) operated this replacement every 100,000 regressions. We increased the threshold to one million regressions, but that still required us to do half a dozen replacements of prior probabilities by estimated posterior probabilities. To check whether this sequential approach affects the results we run one billion random regressions, instead of one million. But we find that the results are almost identical (table 8). In light of this, we are confident our baseline results are not biased by our computational choice.

Another potentially sensitive choice concerns the prior probabilities used to run the BACE approach. Our diffuse prior was that the true model of the economy included one indicator from each of the buckets, and it was reassuring that at least one indicator from eight of the nine buckets was retained as significant. However, it could well be that the posterior probabilities were not independent from our choice of priors. To check whether this is so we assume that the true model of the economy still contains ten explanatory variables, but any of our 31 indicators (30 plus the initial performance of the location) stands the same chance to belong in the true model. Again, the results are highly consistent.

The sample chosen for the analysis could affect the results as well. Our sample included the 21 largest states. To assess whether this choice mattered, we expand the sample to all 31 states for which we have a reasonable amount of data. A challenge in doing so is that two state-level indicators (the labor regulation index and the land reform index) are not available for the additional ten states. To ensure that the comparison is meaningful, we proceed in two steps. First, we exclude these two-state level indicators and re-run the baseline analysis for the largest states. We find that the results from using 29 indicators are highly consistent with the baseline. The estimated impact of some of the indicators changes, but neither the direction nor the significance of the impacts is affected.

As a second step, we expand the sample to all states and conduct the analysis with the available 29 indicators. We then compare the results of using 29 indicators based on all states with the baseline results. The two sets of results are highly consistent. The assessment of the significance of the share of other industries is affected (it becomes positively and significantly related to growth by the sign certainty probability criterion) but the direction and significance of other indicators remains the same as before.

One relevant question is what the BACE approach really adds, relative to a more straightforward OLS regression including all of the potential drivers of growth. The OLS approach is not advisable in a cross-country setting, or when focusing on convergence across states in India, because the number of degrees of freedom rapidly becomes a constraint. But with data available for many hundred districts and several thousand places, this should be less of a concern.

Table 8 Pairwise correlation coefficients between baseline and alternative results

A. Conditional posterior means

	Baseline	1 billion regressions	Alternative prior	29 indicators	All states	OLS, all variables	OLS, 30 indicators
1 billion regressions	1.0000	1					
Alternative prior	0.9967	0.9966	1				
29 indicators	0.9943	0.9941	0.997	1			
All states	0.9804	0.9802	0.9791	0.9774	1		
OLS, all variables	0.9580	0.9585	0.9431	0.9488	0.9352	1	
OLS, 30 indicators	0.9643	0.9646	0.9495	0.9579	0.9447	0.9976	1
Place level regressions	0.9223	0.9225	0.9214	0.9300	0.9460	0.8863	0.8798

B. Posterior inclusion probabilities

	Baseline	1 billion regressions	Alternative prior	29 indicators	All states	OLS, all variables	OLS, 30 indicators
1 billion regressions	0.9998	1					
Alternative prior	0.9805	0.9782	1				
29 indicators	0.9801	0.9783	0.9908	1			
All states	0.9109	0.9093	0.8784	0.8528	1		
Place level regressions	0.5861	0.5953	0.5308	0.5598	0.6154		

To assess whether the use of the BACE approach is justified we run a multivariate OLS regression that includes all 31 indicators. All variables are standardized such that the coefficient estimates are comparable between indicators and with the conditional posterior mean coefficients of the baseline results. Because of the two measures on market assess are not only belong to one bucket but are proxies of the same concept we conduct the same analysis excluding market access based on GDP. On the surface, the difference with the baseline results seems to be minor. Both identify 11 indicators as statistically significant and the correlation between the posterior conditional mean coefficients and the OLS coefficients remains high (albeit not as high as before).

However, there are important differences between the two sets of results. Using the multivariate regression with only one measure on market assess for the comparison, the assessment on railway station density, the share of self-employed, and land regulation index differs substantively from the baseline results. Railway station density becomes one of the least significant indicators according to the OLS estimates, whereas the share of self-employed becomes significant, and the land reform index becomes one of the most relevant indicators. Overall, while the two methods are consistent with each other on many aspect, the BACE approach gives us more confidence because it explicitly addresses uncertainty regarding the true model of the economy.

Finally, the baseline results are based on district-level analysis and a relevant question is whether the conclusions would hold if the analysis was conducted at the place level. One limitation of this comparison is that we can only calculate a few of the indicators at this level of spatial disaggregation. Among them are the indicators for sectoral composition of economic activity, the employment structure, human capital, and social inclusion. For other potential drivers of growth that could well vary across places we can only use the district-level indicators as proxies.

Despite this limitation, the results remain roughly consistent. However, the correlation between posterior inclusion probabilities declines substantially compared to previous robustness tests. Results on the overall significance of the indicators change only for six indicators, but the magnitude of the impacts often differs between district level and place level. This suggests that for a more thorough analysis at the place level, we need to expand the availability of indicators at this level of spatial granularity.

References:

Aiyar, S. (2001). “Growth Theory and Convergence across Indian States: A Panel Study”, in T Callen et al (eds.), *India at the Crossroads: Sustaining Growth and Reducing Poverty*, Washington, DC: International Monetary Fund.

Aghion, P., R. Burgess, D. Redding and F. Zilibotti (2008). “The Unequal Effects of Liberalisation: Evidence from Dismantling the License Raj in India”, *American Economic Review*, 98(4), 1397-1412.

Baddeley, M., K. McNay and R. Cassen (2006). “Divergence in India: Income Differentials at the State Level, 1970-1997”. *Journal of Development Studies*, 42(6), 1000-1022.

Bajpai, N. and J. D. Sachs (1996). “Trends in Interstate Inequalities of Income in India”. *Development Discussion Papers* 528. Cambridge MA: Harvard Institute for International Development.

Bandyopadhyaya, S. (2012). “Convergence Clubs in Incomes across Indian States: Is there Evidence of a Neighbors’ Effect?” *Economics Letters*, 116, 565-570.

Barro, R. J. (1991). “Economic Growth in a Cross Section of Countries”. *Quarterly Journal of Economics*, 106(2), 407-443.

Barro, R., X. Sala-i-Martin, O. J. Blanchard and R. Hall (1991). “Convergence across States and Regions”. *Brookings Papers on Economic Activity*, 107-182.

Barro, R., and X. Sala-i-Martin (1992). “Convergence.” *Journal of Political Economy*, 100(2), 223–251.

Bartelsman, E., J. Haltiwanger and S. Scarpetta (2013). “Cross-country Differences in Productivity: The Role of Allocation and Selection”. *American Economic Review*, 103(1), 305-334.

Besley, T., and R. Burgess (2004). “Can Labor Regulation Hinder Economic Performance? Evidence from India”. *Quarterly Journal of Economics*, 119(1), 91-134.

Besley, T., and R. Burgess (2000). “Land Reform, Poverty Reduction, and Growth: Evidence from India”. *Quarterly Journal of Economics*, 115(2), 389-430.

Baumol, W. (1986). “Productivity Growth, Convergence and Welfare: What the Long-Run Data Show.” *American Economic Review*, 76(5) 1075–1085.

Cashin, P. A., and R. Sahay (1996). “Internal Migration, Centre-State Grants and Economic Growth in the States of India”. *IMF Staff Papers*, 43, 123-71.

Chatterjee, U., R. Murgai, and M. Rama (2015). “Job Opportunities along the Rural-Urban Gradation and Female Labor Force Participation in India”. *Policy Research Working Paper* 7412. Washington DC: The World Bank.

Ciccone, A, and M. Jarocinski (2010). "Determinants of Economic Growth: Will Data Tell?". *American Economic Journal: Macroeconomics*, 2(4), 222-246.

Da Mata, D., U. Deichmann, J. V. Henderson, S. V., Lall and H. G. Wang (2007). “Determinants of City Growth in Brazil”. *Journal of Urban Economics*, 62(2), 252-272.

Das, S. (2012). “The Convergence Debate and Econometric Approaches: Evidence from India” in Ghate, C. (ed.) *The Oxford Handbook of the Indian Economy*. New Delhi: Oxford University Press.

Das, S., C. Ghate and P. E. Robertson (2015). “Remoteness, Urbanization, and India’s Unbalanced Growth”. *World Development*, 66, 572-587.

Datt, G., and M. Ravallion (2002). “Is India's Economic Growth Leaving the Poor Behind?”. *Journal of Economic Perspectives*, 16(3), 89-108.

Datt, G., M. Ravallion, and R. Murgai (2016). “Growth, Urbanization and Poverty Reduction in India”. *NBER Working Paper* 21983. Cambidge MA: National Bureau of Economic Research.

Deaton A, and B. Tten (2017). “Trying to Understand the PPPs in ICP 2011: Why Are the Results so Different?”. *American Economic Journal: Macroeconomics*, 9(1), 243-64.

Duranton, G. (2016). “Determinants of City Growth in Colombia”. *Papers in Regional Science*, 95(1), 101-131.

Duranton, G. (2015). “Growing through Cities in Developing Countries”. *World Bank Research Observer*, 30(1), 39-73.

Fujita, M., P. Krugman, P. R. Venables and M. Fujita (1999). *The Spatial Economy: Cities, Regions and International Trade* (Vol. 213). Cambridge MA: MIT press.

Ghani, E., L. Iyer and S. Mishra (2010). “Are Lagging Regions Catching Up with Leading Regions?” in Ghani, E. (ed.), *The Poor Half Billion in South Asia: What is Holding Back Lagging Regions?*, New Delhi: Oxford University Press.

Ghani, E., W. R. Kerr, and S. O'Connell (2014). “Spatial Determinants of Entrepreneurship in India”. *Regional Studies*, 48(6), 1071-1089.

Ghate, C. and S. Wright (2012). “The ‘V-Factor’: Distribution, Timing and Correlates of the Great Indian Growth Turnaround”. *Journal of Development Economics*, 99, 58-67.

Ghate, C. and S. Wright (2013). “Why Were Some Indian States So Slow to Participate in the Turnaround?”. *Economic & Political Weekly*, 38(13), 118-127.

Glaeser, E. L., J. Scheinkman, and A. Shleifer (1995). “Economic Growth in a Cross-section of Cities”. *Journal of Monetary Economics*, 36(1), 117-143.

Glaeser, E. L., S. P. Kerr, and W. R. Kerr (2015). “Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines”. *Review of Economics and Statistics*, 97(2), 498-520.

Hsieh, C. T., and P. J. Klenow (2009). “Misallocation and Manufacturing TFP in China and India”. *Quarterly Journal of Economics*, 124(4), 1403-1448.

Kalra, S. and P. Sodsriwiboon (2010). “Growth Convergence and Spillovers across Indian States: What Matters? What Does Not?”. *IMF Working Paper* WP/10/96. Washington DC: International Monetary Fund.

Kochhar, K., U. Kumar, R. Rajan, A. Subramanian and I Tokatlidis (2006). “India's Pattern of Development: What Happened, what Follows? ”. *Journal of Monetary Economics*, 53(5), 981-1019.

Krugman, P. R. and A. J. Venables (1995). Globalization and the Inequality of Nations. *Quarterly Journal of Economics*, 110(4):857–880.

Kumar, U., and A. Subramanian (2012). “Growth in India’s States in the First Decade of the 21st Century: Four Facts”. *Economic & Political Weekly*, 47(3), 49.

Kundu, A., and N. Sarangi (2007). “Migration, Employment Status and Poverty: An Analysis across Urban Centres”. *Economic and Political Weekly*, 42(4), 299-306.

Levine, R., and Renelt, D. (1992). “A Sensitivity Analysis of Cross-country Growth Regressions”. *American Economic Review*, 82(4), 942-963.

Li, Y., and M. Rama (2015a). “Firm Dynamics, Productivity Growth, and Job Creation in Developing Countries: The Role of Micro-and Small Enterprises”. *World Bank Research Observer*, 30(1), 3-38.

Li, Y., and M. Rama (2015b). “Households or Locations? Cities, Catchment Areas and Prosperity in India”. *Policy Research Working Paper*, 7473. Washington DC: the World Bank.

Li, Y., M. Rama, V. Galdo, and M. F. Pinto (2015). “A Spatial Database for South Asia”. *Unpublished manuscript*. Washington DC: The World Bank.

Melitz, M. J. (2003). “The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity”. *Econometrica*, 71(6), 1695-1725.

Misra, B. S. (2007). *Regional Growth Dynamics in India in the Post-Economic Reform Period*. London: Palgrave Macmillan.

Ministry of Finance (2017). *Economic Survey*, 2016-17. New Delhi: Government of India.

Nagaraj, R., A. Varoudakis and M.A. Veganzones (2000). “Long-Run Growth Trends and Convergence across Indian States”. *OECD Development Centre Working Paper* 131. Paris: Organisation for Economic Cooperation and Development.

Nayyar, G. (2014). “Economic Growth across Indian States”. *Unpublished manuscript*, Washington DC: The World Bank.

NSSO. (2005). *National Sample Survey of India*, 61st round. New Delhi: National Sample Survey Office, Ministry of Statistics and Programme Implementation, Government of India.

NSSO. (2012). *National Sample Survey of India*, 68th round. New Delhi: National Sample Survey Office, Ministry of Statistics and Programme Implementation, Government of India.

Pritchett, L. (1997). “Divergence, Big Time”. *Journal of Economic Perspectives*, 11(3), 3-17.

Purfield, C. (2006). “Mind the Gap: Is Economic Growth in India Leaving Some States Behind?”. *IMF Working Paper*, 06/103, Washington DC: International Monetary Fund.

Rama, M., T. Béteille, Y. Li, P. K. Mitra, and J. L. Newman (2014). *Addressing Inequality in South Asia*. Washington DC: the World Bank.

Rao, M.G., R.T. Shand and Kalirajan (1999). “Convergence of Incomes across Indian States: A Divergent View”. *Economic & Political Weekly*, 34(13), 769-78.

Redding, S. and A. J. Venables (2004). “Economic Geography and International Inequality”. *Journal of International Economics*, 62(1), 63–82.

Rodrik, D., and A. Subramanian (2005). “From ‘Hindu growth’ to Productivity Surge: The Mystery of the Indian Growth Transition”. *IMF Staff Papers*, 52(2), 193-228.

Romer, P. M. (1986). “Increasing Returns and Long-run Growth”. *Journal of Political Economy*, 94(5), 1002-1037.

Rosenthal, S. S., and W. C. Strange (2004). “Evidence on the Nature and Sources of Agglomeration Economies”. *Handbook of Regional and Urban Economics*, 4, 2119-2171.

Sachs, J.D., N. Bajpai, and A. Ramiah (2002). “Understanding Regional Economic Growth in India”. *Centre for International Development Working Paper*, 88. Cambridge MA: Harvard University.

Sala-i-Martin, X. (1997). “I Just Ran Two Million Regressions”. *American Economic Review*, 87(2), 178-83.

Sala-i-Martin, X., G. Doppelhofer, and R. I. Miller (2004). “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach”. *American Economic Review*, 94(4), 813-835.

Solow, R. M. (1957). “Technical Change and the Aggregate Production Function”. *Review of Economics and Statistics*, 312-320.

Solow, R. M. (1956). “A Contribution to the Theory of Economic Growth”. *Quarterly Journal of Economics*, 70(1), 65-94.

Tewari, M., Godfrey. N., et al. (2016). *Better Cities, Better Growth: India’s Urban Opportunity*. New Climate Economy, World Resources Institute, and Indian Council for Research on International Economic Relations. London, Washington, DC, and New Delhi

Tripathi, S. (2013). “Do Large Agglomerations Lead to Economic Growth? Evidence from Urban India”. *Review of Urban and Regional Development Studies*, 25(3), 176-200.

Tybout, J. (2014). “The Missing Middle, Revisited”. *Journal of Economic Perspectives*, 28(4), 235-36.