

world development report

2009

Reshaping Economic Geography



BACKGROUND PAPER

**URBANIZATION, INEQUALITY AND
ECONOMIC GROWTH: EVIDENCE FROM
INDIAN STATES**

MASSIMILIANO CALI
Overseas Development Institute

Current version: November 2007

Urbanisation, inequality and economic growth: Evidence from Indian states

Preliminary findings of a paper prepared for the WDR 2009

Massimiliano Cali
Overseas Development Institute
November 2007

The aim of this empirical work is to explore various possible implications of the urbanisation process on development outcomes. I investigate these issues in an intra-country rather than in a traditional cross-country setting, using Indian states over the Post-Independence period and Indian towns over the 21st century as the units of analysis. I exploit the richness of contexts within the Indian sub-continent, controlling for many of those countries' unobservables that undermine the robustness of inferences from cross-countries studies.

India has a number of features that make it particularly amenable to this type of empirical verification. First, it is a federal country composed of several states with a fairly high degree of political autonomy, which allows for some state-wise variability in policy variables.

Second, the size of the major states is similar in terms of both population and geographical extension to that of medium-large countries. The average population of the 16 major states considered for the analysis in 2001 was 61,921,484 (Government of India, 2001).¹ If it were a country, it would rank number 20 (between Thailand and United Kingdom) out of 236 (CIA, 2003). Even the least populous state, Jammu & Kashmir with 10,069,917, would rank above the median country (number 70). The average size of the 16 states is 189,573 Km² which would rank number 88 among the largest countries in the world between Senegal and Syria (CIA, 2003). The smallest state is Kerala that with 38,863 Km² would rank number 137 (slightly below the world's median).

The vast size and population of Indian states along with their differences in terms of languages, culture and social norms appear to have limited the mobility of labour across states. Cashin and Sahay (1995) find that the response of migration to income differentials across states was similar to the weak responsiveness of population movements to income differentials across the countries of Europe. Similarly, Topalova (2005) finds extremely limited labour mobility across Indian regions between 1983 and 2000. This relative inter-state immobility of labour is a necessary condition for the empirical test to be meaningful. If that were not the case changes in the labour supply curve in one state may be reflected on urbanisation patterns in other states as well.

Finally, Indian urbanisation experienced an important growth over the Post-Independence period with its rate increasing from 17 percent in 1950 to 27.7% in 2000 (UN, 2006). Lall et al., 2006 estimate that over 20 million people moved from rural to urban areas in the 1990s accounting for 30% of national urban growth. These estimates are consistent with the ones presented below, according to which up to a third of urban population growth over the

¹ The states considered for the analysis are: Andhra Pradesh, Assam, Bihar, Gujarat, Haryana, Jammu & Kashmir, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, West Bengal. Together they represent over 97% of Indian population.

nineties is accounted for by rural-urban migration. As a comparison, migration from rural areas accounted for about 25% of urban growth in the 1980s and 1990s in Africa.

The empirical analysis revolves focuses on three main issues:

- The evolution of rural-urban disparities and its relation with income
- The relation between the speed of urbanisation and income growth
- The evolution of urban concentration

1. Rural-urban disparities

The basic idea of this section is to test whether a relationship exists between urbanisation and rural-urban inequality, and if so what shape it has. I use three families of indices to measure the disparities in welfare between rural and urban areas across Indian states over time:

- 1) Income based measures
- 2) Consumption based measures
- 3) Health based measures

I construct two different rural-urban indicators with the income measures: the difference in *headcount index* between rural and urban areas ($h_1 = h_{rur} - h_{urb}$); and the poverty gap difference ($PG_1 = PG_{rur} - PG_{urb}$). I use the ratio of the rural to the urban mean per capita monthly expenditure as the consumption measure ($ME_2 = ME_{urb} / ME_{rur}$). Unfortunately there is no direct measure of rural and urban access to health services readily available at the state level. I proxy it with the rural-urban difference in death rates per 1000 people ($D_1 = D_{rur} - D_{urb}$). This is a far from ideal indicator of access to health services, as the range of its determinants is likely to be very wide. However, I try to control for some of the main factors (other than access and quality of healthcare) likely to influence this difference. All of these indicators are constructed in such a way that they are increasing in the rural-urban welfare gap.

The basic approach is to estimate the following panel data model:

$$h_{st} = \alpha_s + \gamma_t + \beta_1 y_{st} + \beta_2 y_{st}^2 + \beta_3 (\Delta y_{st} / y_{st-1}) + \delta x_{st} + \varepsilon_{st} \quad (1)$$

where h_{st} is some measure of rural-urban disparities as described above in state s at time t , y_{st} is real income per capita, x_{st} is a vector of socio-demographic controls, α_s is state fixed effects and γ_t is year effects. I estimate it using a fixed effects model. In such a context fixed effects estimation appears to be more appropriate than random effects, as the states considered are very close to the entire population (accounting for over 95% of total Indian population in 2001).² I also run state-level regressions (without controls) to test whether the relationship holds for all states.

² I obtain similar results to those detailed in the main text estimating the model through GLS modelling the error term as an AR(1) process allowing for state-specific autocorrelation.

Data and preliminary results

The data for the income and consumption based measures come from the World Bank dataset prepared by Ozler, Datt and Ravallion (1996), and further updated by the same authors (see the Appendix for a description of the methodology to construct those indices). The same dataset provides also state-wise income data, which have been updated until 2002. Data for death rates come from various years of the Indian Census, and so do the other demographic data.

Table 1 presents the summary statistics for the rural-urban disparity and income variables. Interestingly, not all welfare indicators are worse in rural than urban areas at any point in time. But the average difference in poverty rates between rural and urban areas is 8 percentage points, indicating a substantial, although very variable, gap between rural and urban areas across Indian states.

Table 1: Summary statistics for the main variables

	Head diff	PG diff	Mean ratio	Death diff	Per capita GDP (Rs)	GDP growth
Mean	8.06	2.42	1.39	4.13	12.29	0.05
Std. Dev.	10.79	4.03	0.31	2.04	5.82	0.08
Min	-21.14	-11.03	0.64	-3.90	4.44	-0.27
Max	50.06	14.73	3.08	12.30	39.27	0.43

Figure 1 shows the relationship between rural-urban disparities (using the headcount index) and income per capita for each Indian state over the period 1958-2002. A quite clear non-linear U-shaped relationship emerges, although a few states have the opposite inverted-U shape pattern (i.e. Jammu & Kashmir, Madhya Pradesh and Karnataka) and Orissa, Rajasthan and Tamil Nadu have linear patterns. This U-shaped pattern emerges quite vividly when using the other consumption and health based measures of inequality (Figures 2 and 3). These stylised facts may suggest a pattern of economic development accompanied by a reduction in rural-urban inequalities over time (with an eventual slight increase for certain states). However, if we plot the evolution of GDP per capita and rural-urban disparities over time, the increasing trend appears evident in virtually all states for the former but not for the latter (Figure 4). This calls for a more formal scrutiny of the relationship.

Table 2 presents the results of regressions based on equation (1), which provide support for the U-shaped relation emerging from the graphs. In particular the difference in the headcount index decreases as income rises up to a point after which it starts increasing again. In particular in the baseline regression (column 1) a 10% increase in real per capita GDP is associated with a reduction of 0.5 percentage points in rural-urban difference in the headcount index. The trough in this difference is reached for a value of real GDP per capita of 26.2 Rs. after which the difference starts rising. However, only 3 out of 16 states had income per capita higher than this level in 2000. Rural-urban inequality increases in the speed of income growth, although the coefficient is significant only at the 10% level. These results are robust to the inclusion of socio-demographic controls (column 2), with the share of the population in working age in urban (rural) areas being positively (negatively) associated with rural-urban difference. The opposite is true for the share of population over 60. The female/male ratio in rural areas has a negative effect on the difference, although it is not significant at conventional

levels, and so is the log of total population.³ In column 3 it appears that the Indian trade liberalisation of 1991 is associated with an increase in rural-urban inequalities, although the inequality reducing impact of income growth is accentuated after this period.⁴

Table 2: Rural-urban disparities and income per capita across Indian states, 1958-2002

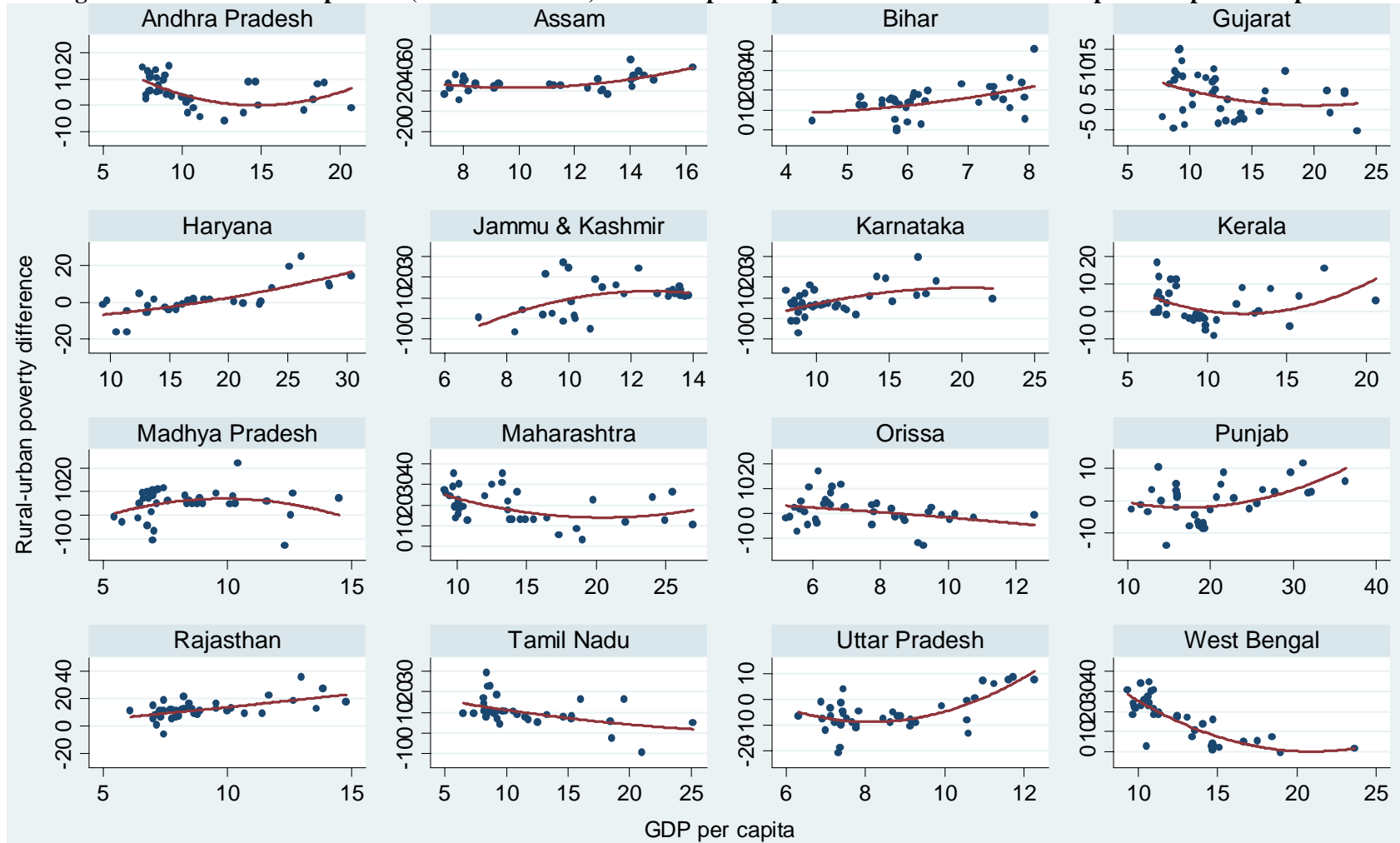
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Head diff	Head diff	Head diff	PG diff	Mean ratio	Death diff.	Death diff.	Death diff.
GDP pc	-2.465 (4.54)**	-2.090 (3.75)**	-1.382 (2.65)**	-0.567 (1.81)	-0.070 (3.18)**	-0.698 (11.60)**	0.109 (1.59)	-0.115 (1.35)
GDP sq.	0.047 (4.37)**	0.038 (3.73)**	0.042 (3.24)**	0.016 (2.58)*	0.001 (2.80)**	0.013 (8.01)**	0.001 (0.64)	0.004 (2.59)*
GDP growth	6.820 (1.76)	8.733 (2.44)*	5.406 (1.76)	2.902 (1.73)	0.165 (1.52)	2.805 (4.04)**	0.379 (0.53)	1.125 (1.48)
Rural 15-59		-0.549 (1.57)	-1.306 (3.68)**	-0.613 (4.40)**	0.004 (0.36)			
Urban 15-59		0.561 (3.65)**	0.418 (2.78)**	0.164 (2.12)*	0.013 (2.89)**			
Rural 60+		2.362 (1.70)	1.499 (1.06)	1.311 (2.01)*	0.019 (0.45)			0.608 (2.68)**
Urban 60+		-5.287 (2.90)**	-3.674 (2.18)*	-1.336 (1.45)	-0.103 (1.92)			-1.131 (4.20)**
Fem/male (15-34 rur)		-30.806 (1.10)	-8.449 (0.29)	3.363 (0.25)	1.874 (1.88)			-13.448 (2.92)**
Ln pop.		1.023 (0.08)	1.105 (0.28)	-7.112 (1.05)	-0.485 (1.19)			-1.593 (0.70)
Rural 0-14								-0.430 (6.31)**
Urban 0-14								-0.048 (1.90)
GDP*1992			-0.414 (1.67)					
1992			16.504 (4.14)**					
Year effects	YES	YES	NO	YES	YES	NO	YES	YES
State effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	564	522	522	462	462	448	448	403
R-squared	0.64	0.71	0.68	0.68	0.79	0.64	0.81	0.85

*Robust t-statistics in parenthesis; * significant at the 5% level; ** significant at the 1% level.*

³ Note that due to data availability the rural female/male ratio is referred to the population in the cohort 15-34 years of age. This share is likely to be a good proxy for the female/male ratio in total rural population.

⁴ Note that this result is obtained without the inclusion of year effects, thus it could only be signalling a generalised increase in rural-urban inequalities (i.e. urban poverty being reduced faster than rural poverty), which has occurred in the last two decades in Indian states. However, the results are not as neat when I run the same regressions using earlier years (i.e. 1988, 1989 and 1990) as break points (not shown here).

Figure 1: Rural-urban disparities (headcount index) and GDP per capita across Indian states in the post-independence period



Note: Poverty difference is measured as the difference between the poverty headcount index in rural areas and that in urban areas

Figure 2: Rural-urban disparities (Mean consumption) and GDP per capita across Indian states, 1958-1994

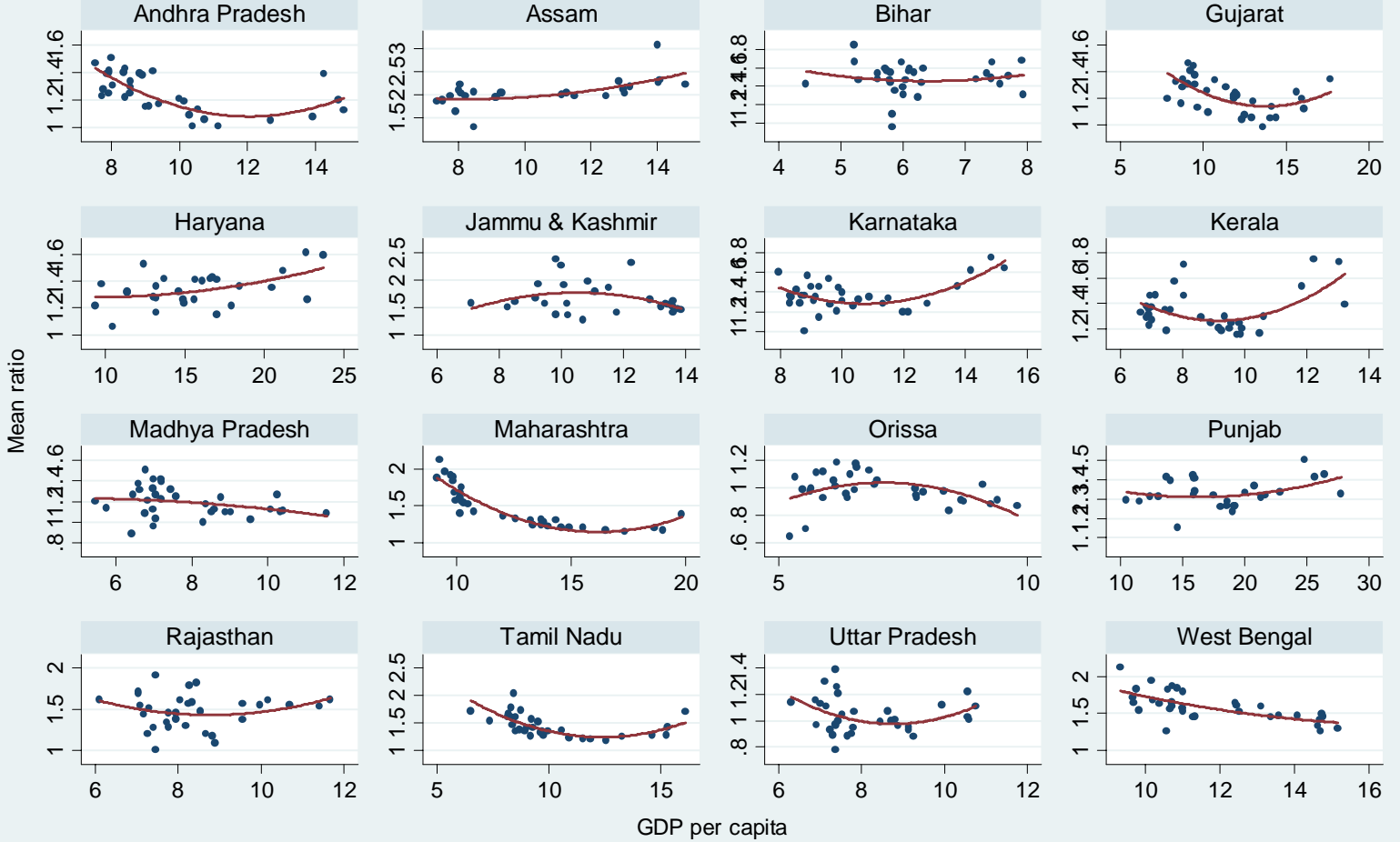


Figure 3: Rural-urban death rates disparities and GDP per capita across Indian states, 1971-2001

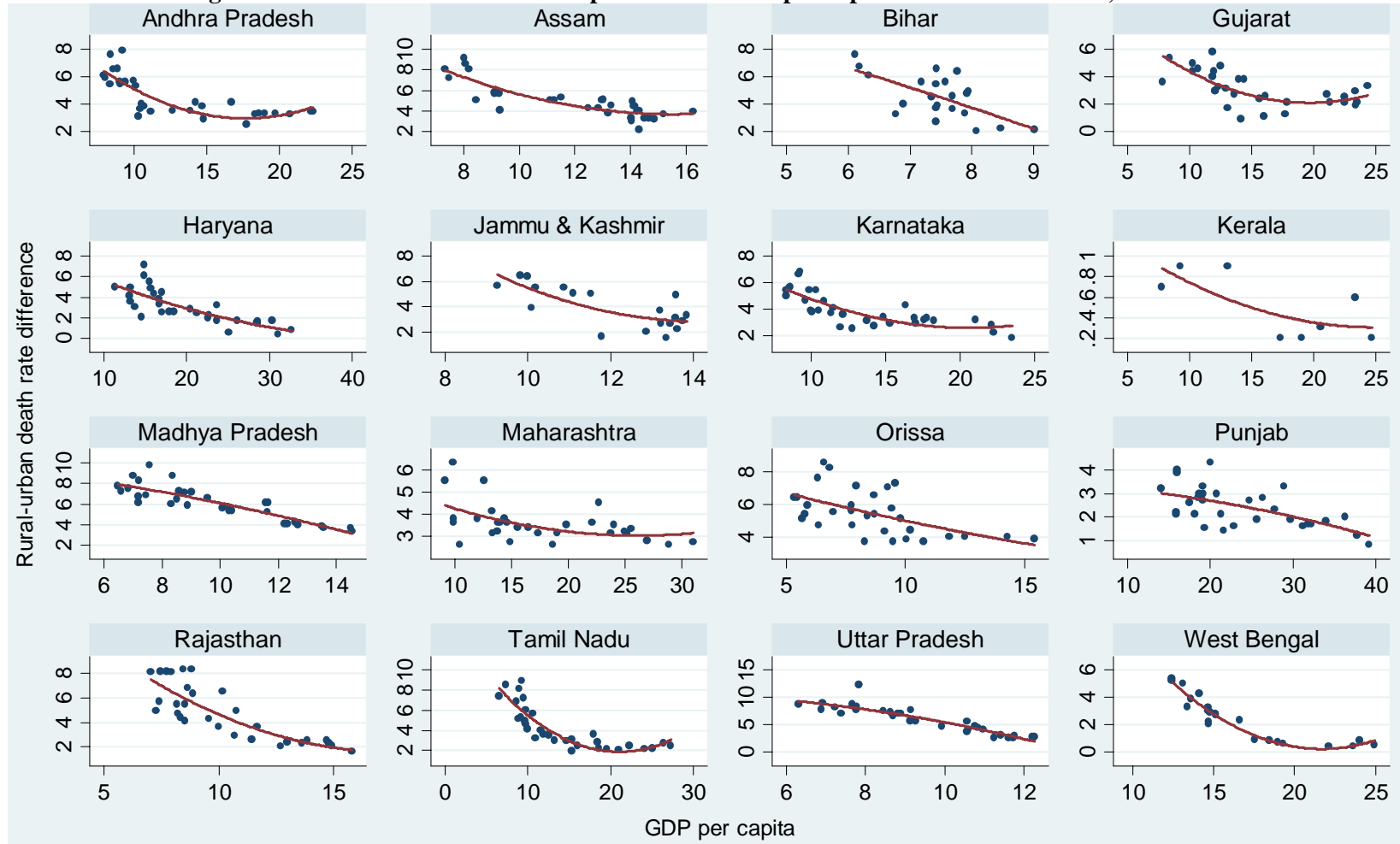
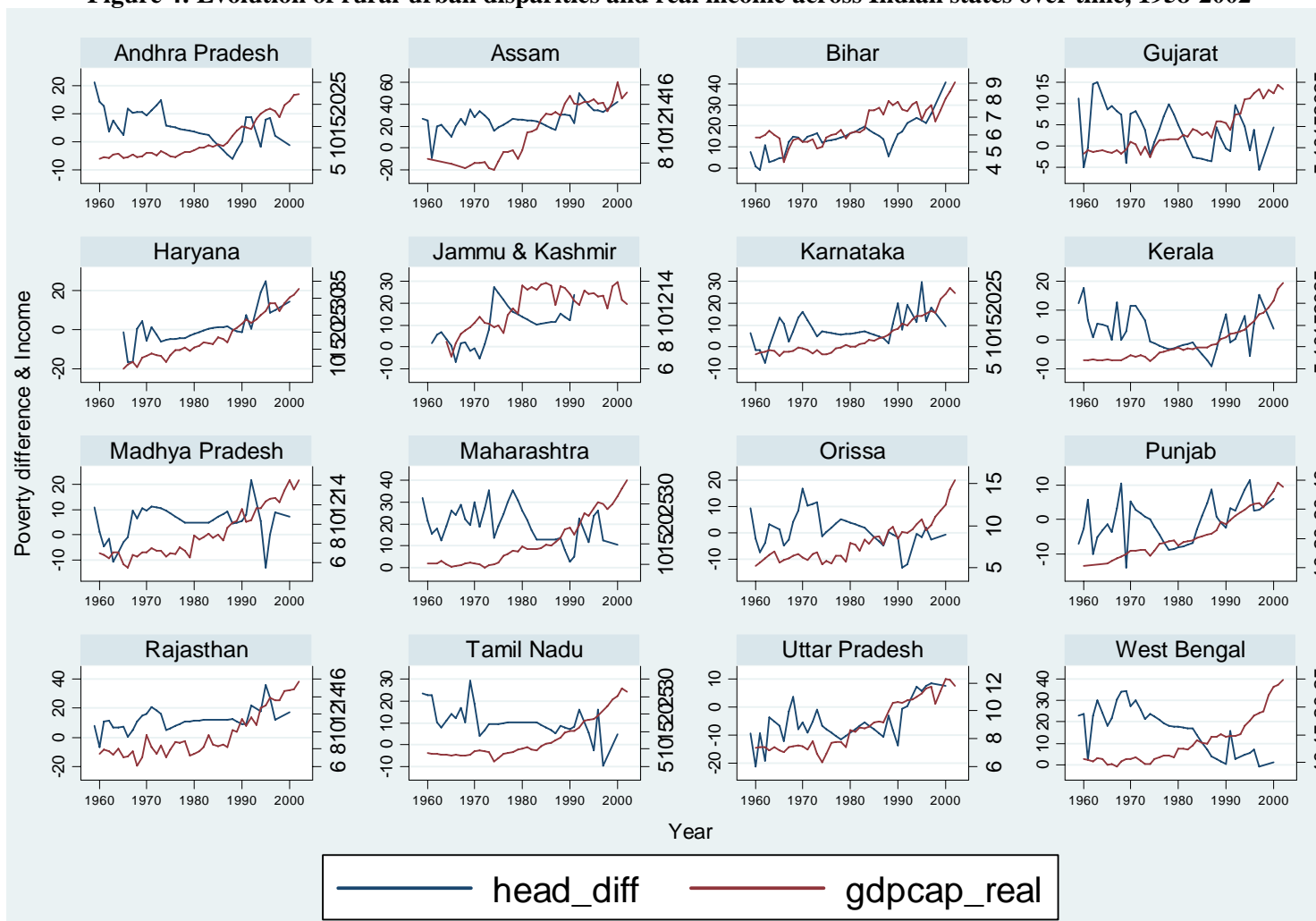


Figure 4: Evolution of rural-urban disparities and real income across Indian states over time, 1958-2002



These results are robust to the use of the other measures of rural-urban disparities, i.e. the difference in poverty gap and the ratio of real mean consumptions (columns 4 and 5). In the latter regression the age composition controls lose some significance, while the female/male ratio in rural areas is positively and significantly (at the 10% level) associated with an increase in the gap in mean consumption (urban over rural), suggesting an anti-female bias in consumption across Indian states.

When I test the disparity-income relation using the death rate as dependent variable, the results are consistent with the previous ones in the model without year effects (column 6), while the introduction of year effects reverse the sign of income (although not statistically significant) – column 7.⁵ This could be the effect of omitted variable bias, which may be particularly relevant in the case of death rates, as other factors may be crucially determining death rates. As a matter of fact the U-shaped relation between disparities and income emerges again once the age and gender composition variables are included (with the share of elderly - over 59 – increasing the death rate and the share of young – below 15 – reducing it). Also, *ceteris paribus* a higher share of female in the rural population reduces the disparities in death rates. This can be related to females having lower mortality rates than males and possibly helping to improve the access of the population to healthcare services (e.g. women in general devote higher parts of the family budget to social spending than men).

2. Urbanisation and growth

I next investigate how the process of urbanisation is linked to that of economic development. As in the case of income and rural-urban inequality, it is appropriate to think about this link as a structural correlation rather than a causal relationship. The basic specification to test this correlation is similar to the one in (1), and is defined as:

$$\ln(y)_{st} = \alpha_s + \gamma_t + \beta_1 \ln(y)_{st-1} + \beta_2 (\Delta u_{st} / u_{st-1}) + \delta x_{st} + \varepsilon_{st} \quad (2)$$

where u_{st} is the urban population of state s at time t . In this way I try to capture the relationship between the growth in income and the proportionate growth in the urban population. I estimate it both by fixed effects and by GLS with state-specific disturbances modelled as an AR(1) process.

Data and preliminary results

The urban population data come from various publications of the Indian Census (between 1951 and 2001) and have a ten-year frequency, thus the number of observations is limited relative to the other analysis.

Table 3 presents the results of the estimation of (2). By and large the results suggest that urban growth tends to be negatively related to income growth, although this finding is not statistically very robust especially when other controls are included. This is evident in column 2, where the low significance of β_2 in the basic specification (column 1) is further reduced by the addition of demographic controls (i.e. the share

⁵ The results are different for random effects estimation, but the Hausman test of random vs. fixed effects estimator, strongly rejects the null of no systematic difference between coefficients estimated using the two methods. Therefore RE estimation may yield biased coefficients.

of population above 59 in total population and its square term, the log of total population and the female/male ratio as described above). These results are stronger when I estimate (1) through GLS AR(1), as in column 3, where the negative relationship between income growth and urban growth is significant at the 5% level.

When I use proportionate income growth ($\Delta y_t/y_{t-1}$) as the left-hand side variable, the intensity of the relationship with urban growth is the same as with GDP per capita for the FE model (column 3) and it is only slightly reduced in the GLS model (column 5). Further, there is some evidence of mean reversal in income, i.e. higher income in one period is associated with lower growth in the subsequent one.

Next I test the correlation between the share of urban population and growth using log of urban population and log of total population as explanatory variables. For a given of urban population, the total population tends to be negatively associated with GDP growth, indicating a positive relation between the level of urbanisation and growth. However, the negative sign of the coefficient of urban population, although insignificant, reduces this effect.⁶ And this positive association disappears altogether when the demographic controls are added (column 7).

Finally, I address the issue of whether urbanisation is positively associated with the level of economic development. In the last two columns I regress real GDP per capita on the share of urban population, finding a positive and significant relationship, as expected. However, the significance of the relation drops below standard levels once the set of controls is added (column 9). This suggests that the relation between urbanisation and income per capita is not a strong one when we consider it within an individual state over time. This relation is much stronger across states, as it emerges from the results of the regressions without state fixed effects (not reported here).

The level of urbanisation and that of economic development seem to go hand in hand within Indian states over time, but this relation does not appear to be a very robust one. On the other hand, it emerges quite clearly that the rate of urbanisation (i.e. how fast a state urbanises) and the rate of growth are negatively correlated (if anything). This finding is somewhat surprising: in the ‘average’ Indian state, periods of faster urbanisation tend to be associated with periods of slower growth. Whether such a result points towards an urbanisation driven by push rather than pull factors could be interesting matter of further research.

⁶ In line with this finding, when I use urban share as the main regressor instead of log of urban and total population, its coefficient is positive but not significant.

Table 3: Economic growth and urban growth across Indian states, 1961-2001

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Ln GDP	Ln GDP	Ln GDP	Δ GDP	Δ GDP	Ln GDP	Ln GDP	Ln GDP	Ln GDP
	FE	FE	GLS AR(1)	FE	GLS AR(1)	FE	FE	FE	FE
Ln(GDP _t)	0.986 (14.40)**	0.765 (9.24)**	0.958 (59.57)**	-0.242 (2.81)**	-0.036 (1.98)*	0.920 (12.56)**	0.748 (9.01)**		
Urban growth	-1.261 (1.38)	-0.895 (1.09)	-0.963 (2.27)*	-0.876 (1.03)	-0.861 (1.66)				
Ln urban pop						-0.075 (0.70)	-0.070 (0.73)		
Urban share								2.418 (2.76)**	0.965 (1.35)
Share 60+		-0.154 (1.43)	-0.176 (3.56)**	-0.169 (1.50)	-0.113 (1.84)		-0.156 (1.43)		-0.082 (0.46)
Share 60+ sq.		0.015 (2.05)*	0.014 (3.89)**	0.016 (2.10)*	0.010 (2.16)*		0.015 (2.05)*		0.014 (1.19)
Ln tot pop		0.049 (0.31)	-0.025 (3.02)**	0.051 (0.30)	-0.029 (3.18)**	-0.303 (1.97)	0.028 (0.18)		0.011 (0.04)
Fem/male (15-34 rural)		-0.197 (0.52)	-0.172 (2.08)*	-0.251 (0.64)	-0.087 (0.98)		-0.164 (0.43)		0.718 (1.16)
1971	-0.025 (0.98)	-0.038 (0.90)	-0.019 (1.55)	-0.035 (0.81)	-0.019 (1.17)	0.062 (1.21)	-0.014 (0.26)	0.013 (0.26)	0.008 (0.11)
1981	0.014 (0.46)	0.004 (0.05)	0.024 (1.58)	0.008 (0.10)	0.022 (1.11)	0.199 (1.99)	0.057 (0.58)	0.096 (1.45)	0.092 (0.64)
1991	-0.024 (0.57)	0.001 (0.01)	-0.002 (0.11)	0.005 (0.04)	0.006 (0.27)	0.276 (1.83)	0.093 (0.64)	0.286 (3.46)**	0.290 (1.37)
2001	0.017 (0.28)	0.087 (0.56)	0.042 (1.89)	0.092 (0.56)	0.049 (1.96)*	0.415 (2.11)*	0.213 (1.11)	0.567 (5.46)**	0.604 (2.24)*
Observations	76	70	70	70	70	76	70	76	70
Number of state	16	15	15	15	15	16	15	16	15
R-squared	0.97	0.98		0.43		0.98	0.98	0.89	0.95

Robust t-statistics in parenthesis; * significant at the 5% level, ** significant at the 1% level; Jammu is the state excluded in the columns with demographic variables, as data are not available.

3. The evolution of urban areas: is there a tendency towards concentration?

In order to analyse the determinants of the growth of Indian cities over time, I compiled a dataset of Indian towns and urban agglomerations' population for the 20th century (from the Indian Census). It is important to understand the distinction that the Census makes between towns and urban agglomerations (UAs), as this will feature prominently in the analysis. UAs are groups of towns that belong to the same urban area as defined it by the Census. These agglomerations usually comprise a core large town, surrounded by a number of smaller towns. Sometimes the difference in population between the main town and the UA may be substantial. For example the Calcutta UA had a population of 13.2 million in 2001 while the town of Calcutta had 4.6 million. The former comprises over 100 towns, many of which have been incorporated to the UA over time. The incorporation of a new town may bias the analysis as it provides a source of growth which is lumpy and has little relation with socio-economic characteristics of the UA. Therefore throughout the analysis I try to focus on towns (which are not subject to this problem), while analysing separately UAs.

All towns and urban agglomerations with a population over 10,000 in 1991 are included, while the coverage for urban areas below 10,000 is patchy. This translates into an average of almost 2,500 observations per period for a total of 11 periods (i.e. 1901-2001 with a ten-yearly frequency). Table 4 provides summary statistics for the two main variables used in the analysis: population and ten-year population growth rate. The latter is computed using the formula: $g = (u_t / u_{t-10})^{1/10} - 1$ where u_t is population at time t. Both the number of towns and urban agglomerations and their average population increase over the century following India's urbanisation process. Interestingly, the process intensifies over time, as it is indicated by the increase in the mean growth rate of urban areas over the 20th century, at least until 1981, after which there is a slight drop in growth rate.

Table 4: Summary statistics for cities' population and population growth, 1901-2001

	Population			Growth rate		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
1901	1,445	22,661	67,757			
1911	1,487	23,148	76,908	1,390	-0.17%	2.93%
1921	1,580	23,357	80,792	1,462	0.38%	2.53%
1931	1,716	25,739	86,651	1,561	1.41%	2.04%
1941	1,893	31,975	128,275	1,703	1.88%	2.32%
1951	2,213	39,025	173,995	1,861	2.21%	3.00%
1961	2,382	48,829	222,716	2,003	2.49%	2.91%
1971	2,762	59,226	284,855	2,369	2.86%	2.54%
1981	3,294	71,098	357,531	2,714	3.27%	2.27%
1991^a	4,428	72,771	402,377	3,287	2.75%	2.56%
2001^b	3,943	96,539	544,474	3,936	2.17%	2.17%

a. The mean value for 1991 is not strictly comparable to that of the other years due to the wider cities' coverage; b. the distribution of town for the year 2000 is slightly skewed towards larger towns due to data availability.

The analysis concentrates on testing whether the Indian urban system has evolved towards more or less concentration during the 21st century. I am going to use two

main specifications to test for concentration. The first focuses on the effect of city size of its subsequent rate of population growth:

$$g_{it} = \alpha_i + \gamma_t + \beta_1 \ln(u)_{it-10} + \beta_2 [\ln(u)_{it-10}]^2 + \varepsilon_{it} \quad (3)$$

where g_{it} is the annual rate of growth (as defined above) of town (or UA) i at time t . The test for concentration is captured by β_1 and β_2 coefficients. In particular a negative sign on the former indicates a tendency towards concentration (i.e. larger towns grow faster), while the latter coefficient captures eventual non-linearity in the relation. I complement this analysis which tests for the population-growth relation at the town level, with one focusing on groups of towns, which are aggregated in classes according to their size:

$$g_{it} = \alpha_s + \gamma_t + \mathbf{B} \sum_{j=1}^5 c_{jt-10}^i + \varepsilon_{st} \quad (4)$$

where c_{jt-10}^i takes the value of one if the town (or UA) i belongs to the class j at time $t-10$ and zero otherwise, and α_s are state effects. The Census identifies six classes of towns that I find useful to re-aggregate for two reasons. First, the Census classes are based on the availability of data for all towns, so the lowest two classes cover towns below 10,000 (while my data cover mainly towns above 10,000 in 1991). Second, these classes do not provide an accurate representation of the upper tail of the town distribution, which is all lumped in class I (including all towns above 100,000). I find it useful to re-aggregate these classes as follows: class VI: below 19,999; class V: 20,000-49,999; class IV: 50,000-99,999; class III: 100,000-299,999; class II: 300,000-999,999; class I: above 1,000,000. I also change this classification to test its robustness to subjective division.⁷

The results from specification (3) are presented in Table 5 and yield support for the idea of de-concentration over time. City size exerts a significant negative influence on subsequent growth, although the intensity of the effect diminishes with size. In particular, a 1% increase in population determines a reduction 0.15% in the average annual rate of growth of (from the average value of 2.1% - column 1). These results are valid for both the pre- and pos-Independence periods (columns 2-3), as well as when considering only UAs (column 4), only towns belonging to UAs (column 5) and only the large UAs (column 6). In the last three cases the inverse size-growth relationship appears to be even more marked than in the baseline case. The non-linearity seems to fade for very large towns (over 500,000) – column 7. In this case the negative effect appears to be linear.⁸

The main interpretation of these results is that as a town (or UA) grows in size, its rate of growth slows down relative to the rate of growth experienced when its size was smaller. This result is statistically more important than the cross-sectional one, i.e. larger towns grow more slowly than smaller ones, as it is evident from two facts. First, the FE regressions (in the first seven columns) explain a much larger part (by over 100 times) of the within group than the between groups variation; second the

⁷ In particular, I also use another classification which gives a more balanced allocation of towns across classes: class VI: below 9,999; class V: 10,000-29,999; class IV: 30,000-79,999; class III: 80,000-199,000; class II: 200,000-699,999; class I: above 700,000.

⁸ I regress the growth rate on the linear term, which is significant at the 1% level (not reported here).

intensity of the size effect reduces significantly in the OLS relative to the FE estimation, as shown in column 8. However, the estimation without town effects confirms the validity of the U-shaped relationship between growth and size, even when I include district effects to control for local conditions likely to influence urban growth (column 9).

Table 5: Convergence in growth rates across Indian cities in the 20th century?

	(1)	(2)	(3)	(4)	(5)	(6)
Periods	All	Post-Ind	Pre-Ind	All	All	All
Model	All towns (FE)	All towns (FE)	All towns (FE)	UA only (FE)	UA outr. (FE)	Large UA (FE)
Ln(pop) ₋₁₀	-0.075 (11.99)**	-0.078 (9.36)**	-0.196 (6.36)**	-0.080 (5.79)**	-0.093 (8.65)**	-0.079 (4.63)**
Ln(pop) ₋₁₀ sq.	0.002 (7.69)**	0.001 (3.77)**	0.005 (3.28)**	0.002 (3.72)**	0.003 (6.29)**	0.001 (2.40)*
Year Effects	YES	YES	YES	YES	YES	YES
Observations	20526	15040	3949	2887	3793	1760
No. of groups	4207	4207	1429	374	876	237
R-squared	0.31	0.35	0.61	0.33	0.31	0.34
	(7)	(8)	(9)	(10)	(11)	
Periods	All	All	All	1991-01	1981-91	
Model	Towns > ½ mln (FE)	All towns (OLS)	All towns dist. effects	All towns dist. effects	All towns dist. effects	
Ln(pop) ₋₁₀	-0.069 (1.53)	-0.033 (6.91)**	-0.031 (5.33)**	-0.021 (2.28)*	-0.024 (1.19)	
Ln(pop) ₋₁₀ sq.	0.001 (1.00)	0.0015 (6.87)**	0.0015 (5.27)**	0.001 (2.33)*	0.001 (1.11)	
Ln(dist_cap)				-0.0067 (4.46)**	-0.0077 (3.21)**	
River				-0.004 (3.51)**	-0.003 (2.50)*	
Part of UA					0.004 (2.20)*	
Year Effects	YES	YES	YES	NO	NO	
Observations	205	20526	20517	3818	3029	
No. of groups	69	0	474	450	447	
R-squared	0.44	0.14	0.19	0.21	0.30	

*Robust t-statistics in parenthesis; * significant at the 5% level, ** significant at the 1% level; dependent variable: annual growth rate in population.*

This relationship is less evident at the cross-sectional level. It is valid but significant only at the 5% level, when I regress the 1991-2001 annual growth rate on towns' size (including district effects and geographical controls) – column 10. And the significance of the β coefficients disappears when I consider the 1991-2001 growth rate (column 11), although the signs are the same. The results from the last two columns suggest that the strength of the growth convergence effect across towns may not be as significant as that over time. The cross-sectional analysis further highlights that the distance from the state capital negatively affects the town's growth prospects, and so does the presence of a navigable river.⁹ On the other hand towns part of UAs

⁹ This is a dummy variable which takes the value of 1 if the town has a navigable river.

(proxy for being close to large centers) enhances its growth prospects. Geographical location does seem to matter for cities' growth in India, although further analysis would be needed to draw more robust conclusions.

Table 6: Class city size and population growth across size classes, 1901-2001

Periods	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model	All	All	All	All	2001	1991	1981	Pre-Ind
	Town	UA	Different	State	State	State	State	State
	effects	effects	classif.	effects	effects	effects	effects	effects
			(FE)					
Class 1	-0.017 (4.99)**	-0.026 (4.99)**	-0.021 (7.68)**	0.000 (0.05)	0.004 (1.19)	-0.008 (1.71)	-0.001 (0.11)	0.003 (0.33)
Class 2	-0.008 (3.91)**	-0.016 (4.36)**	-0.008 (4.83)**	0.003 (2.29)*	0.005 (1.89)	0.006 (2.23)*	-0.002 (0.73)	0.005 (0.94)
Class 4	0.008 (6.15)**	0.012 (5.14)**	0.014 (10.26)**	-0.002 (2.28)*	-0.002 (1.31)	-0.002 (1.42)	-0.001 (0.63)	-0.005 (1.92)
Class 5	0.018 (11.37)**	0.027 (6.20)**	0.029 (14.86)**	-0.002 (2.15)*	-0.002 (1.57)	-0.000 (0.22)	-0.002 (1.03)	-0.004 (1.93)
Class 6	0.030 (14.31)**	0.043 (6.30)**	0.043 (15.95)**	-0.001 (1.15)	-0.002 (1.12)	0.002 (1.59)	-0.000 (0.16)	-0.005 (2.81)**
Year effects	YES	YES	YES	YES	NO	NO	NO	YES
Observations	20526	2887	20526	20526	3845	3061	2490	5486
No of groups	4207	374	4207	31	27	31	30	23
R-squared	0.16	0.17	0.18	0.14	0.03	0.09	0.06	0.13

*Robust t-statistics in parenthesis; * significant at the 5% level, ** significant at the 1% level; dependent variable: annual growth rate in population.*

The analysis using the classes of cities instead of the initial population confirm the tendency of towns to slow down their growth as they become larger (Table 6). When a town becomes Class I, it grows less than when it was Class II, and so on (column 1). This is the case also for UAs (column 2), and the results are robust to using a different classification of towns as described above (column 3). Things change when I include only state effects but not town effects (column 4). A town in class II (medium-large sized) is more likely to grow than any other town (controlling for the state), while a town in class I (medium-small) is likely to grow the least. Towns in class V (small towns) also grow slower, while the growth rates for the other classes are not statistically different from towns in class III. These broad results hold fairly well when considering only growth in 1991-2001 (column 5), but in the 1981-1991 period class I towns have been the ones experiencing the lowest growth (column 6). The results are unclear for the 1971-81 period (column 7) while they indicate a bias against small towns in the pre-Independence period – column 8 (although this may just be the product of a classification which is less meaningful for a period in which most towns would be classified in class VI and V). The interpretation of these results require further scrutiny.

Rural-urban migration

How much of cities' growth is actually generated by rural-urban migration. Table 7 tries to address this question by detailing the share of rural-urban migrant (in the previous ten years) in total rural and urban population. For example 11.6% of Andhra

Pradesh urban population has moved from rural to urban areas in the previous ten years. The number of rural-urban migrants for each state has been calculated using data on urban population, birth and death rates (and assuming inter-state immobility of labour) in the following way:

$$M_{st} = U_{st} - \sum_{i=t-9}^t [(bu_{si} - du_{si})U_{si}]$$

where U_{st} is total urban population in state s at time t , bu_{st} and du_{st} are birth rate and death rate of the urban population respectively.

For most states this share decreases between 1981 and 1991 (with Jammu & Kashmir, Orissa and Assam having the highest shares in 1991). In 2001 this trend is reversed in half of the states, suggesting that the opening up of the Indian economy in the early nineties may have spurred some internal migration from rural to urban areas. However, the data is patchy and the evidence is not robust enough to make more rigorous inferences for the time being.

Table 7: share of rural-urban migrants in the population across Indian states, 1981-2001

	Share of urban pop			Share of rural pop		
	1981	1991	2001	1981	1991	2001
Andhra Pradesh	11.6%	11.6%	4.6%	3.5%	4.3%	1.7%
Assam	14.0%	12.2%	12.4%	1.5%	1.5%	1.8%
Bihar		4.4%	5.4%		0.7%	0.8%
Gujarat	9.0%	7.2%	7.8%	4.1%	3.8%	4.7%
Haryana	14.8%	11.9%	13.0%	4.2%	3.9%	5.3%
Jammu & Kashmir	17.2%	14.8%		4.6%	4.5%	
Karnataka	14.4%	9.6%	6.5%	5.8%	4.3%	3.4%
Kerala	7.2%	11.1%	19.2%	1.7%	4.0%	6.7%
Madhya Pradesh	14.7%	11.9%	9.7%	3.7%	3.6%	4.9%
Maharashtra	11.1%	10.0%	10.0%	6.0%	6.3%	7.4%
Orissa	21.0%	13.8%	9.4%	2.8%	2.1%	1.6%
Punjab	10.2%	6.8%	8.8%	3.9%	2.9%	4.5%
Rajasthan	16.2%	10.8%	6.8%	4.3%	3.2%	2.1%
Tamil Nadu	7.3%	3.2%	10.9%	3.6%	1.6%	8.5%
Uttar Pradesh	16.3%	11.9%	6.8%	3.6%	2.9%	1.8%
West Bengal		10.2%	8.7%		3.9%	3.4%

References

- Cashin P. and R. Sahay, 1996. Internal Migration, Center-State Grants and Economic Growth in the States of India, IMF Working Paper 95/58.
- CIA, 2003. The World Factbook 2002.
- Datt, G., 1995. Poverty in India 1951-1992: Trends and Decompositions, Policy Research Department, World Bank, Washington, mimeo.
- Government of India, various years. *Statistical Census of India*, 1951, 1961, 1971, 1981, 1991, 2001.
- Kuznets, S., 1955, Economic Growth and Income Inequality, *American Economic Review*, 1, Vol. XLV.
- Lall S.V., H. Selod, and Z. Shalizi, 2006. Rural-urban migration in developing countries: A survey of theoretical predictions and empirical findings. Policy Research Working Paper 3915, World Bank.
- Ozler, Datt and Ravallion, 1996. *A Database on Poverty and Growth in India*, Washington: World Bank.
- Planning Commission, 1993. Report on the Expert Group on the Estimation of the Proportion and Number of Poor, Delhi: Government of India.
- Topalova, P., 2005. Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts, NBER Working Paper No. W11614.
- United Nations, 2006. World Urbanization Prospects: The 2005 Revision Population Database, <http://esa.un.org/unpp/>.

Appendix 1

Methodological note to the construction of poverty measures

Headcount Index

The headcount index measures the share of the population below the poverty line in a certain geographical unit. The poverty lines used by the dataset are those recommended by the Planning Commission (1993) and are as follows. The rural poverty line is given by a per capita monthly expenditure of Rs. 49 at October 1973-June 1974 all-India rural prices. The urban poverty line is given by a per capita monthly expenditure of Rs. 57 at October 1973-June 1974 all-India urban prices (see Datt (1995) for further details on the rural and urban cost of living indices and the estimation of poverty measures).

Poverty Gap Index

This is calculated as the mean distance below the poverty line as a proportion of the poverty line where the mean is taken over the whole population, counting the non-poor as having zero poverty gap. That is the mean shortfall from the poverty line (counting the non poor as having zero shortfall), expressed as a percentage of the poverty line.