INDUSTRIAL DECENTRALIZATION AND URBAN DEVELOPMENT in India with Consideration of South-East and East Asian Cases

A Workshop on a MOST/UNESCO Research Project

Foreword
François HOULLIER

Véronique BÉNÉÏ & Loraine KENNEDY
Editors

23

Pondy Paper in Social Sciences

INSTITUT FRANÇAIS DE PONDICHÉRY

Département de Sciences Sociales
BP 33, Pondichéry, 605 001
Composition - Cellule d'édition IFP
Impression - All India Press, Pondichéry

CHARACTERISTICS AND CORRELATES OF URBANIZATION IN INDIA

An illustration of the use of the urban database

Christophe Z. GUILMOTO

This paper should serve to illustrate two applications of the database of Indian towns. We have prepared various tables, models and graphs about Indian urbanization from these census data. The first series of tables and graphs provides descriptive statistics of the urban Indian universe. From the base, specific data can be extracted and cities identified according to research needs. It may be useful to stress that most descriptive statistics could likewise be mapped on a regional or national basis though we have not included any map here.

A further series of statistical results is based on more sophisticated processing and modelling. Here, we have tried to give a few examples of the analyses that can be performed from our information base. For both the query and analysis applications of the base, the number and diversity of possible investigations remains of course bigger than the few illustrations we provide in this paper.

1. PRESENTATION OF SOME ECONOMIC AND SOCIAL CHARACTERISTICS OF INDIAN TOWNS

To begin with, we will present the result of simple tabulations of some of the variables at our disposal. We are using here the whole urban sample, without differentiating between towns according to their location, size or any other characteristics. In the analyses that follow, the results have not been weighted by population size and large cities therefore carry exactly the same weight as smaller units, some of which with population less than 5,000.

Table 1 indicates the average value of each variable over the entire urban sample, as well as the number of towns with more prominent values. As can be seen, the Harijan component of urban populations is quite important, and stable, which is not the case with the tribal population. Sex ratio (measured as the number of women per 1,000 men) is significantly below the national average because of the

impact of sex-selective migrations in urbanization. To some extent, sex ratio can even be used as a proxy to migration intensity, though there are deep regional variations (towns in South India are much more "feminine"). The child per woman ratio is a rough indicator of urban fertility level.

The last variables described on Table 4 refer to population density at a residence, household or town level. Population per household is closely related to both family structure and couple fertility. For reasons mentioned by J.-F. Sevoz, town density is a very unreliable measurement.

Tables 5 and 6 present the economic distribution of the urban workforce. Table 5 shows the variation between male and female participation rates and economic activities. We have underlined some of the most significant variations related to the relative importance of women among agricultural labourers and other services. Table 6 indicates that the agricultural component of the urban workforce is quite sizeable, as it is more than half of the total in several hundreds of (small) towns.

The following tables are derived from cross-tabulations by urban size. As the size of urban settlements is one the most meaningful dimensions of urbanization in the MOST project, we have selected this variable to provide some further examples of data processing and modelling. The size of urban population is usually expressed in terms of town class (I to VI) according to the standard Census of India definition. In other illustrations, the town size is given either in absolute values or in log values.

Tables 4-5 and Figures 1-4 show the close relationship between population size and a large number of social and economic variables such as agricultural activities, literacy or housing amenities. There is therefore a somewhat central component of urbanization which involves a large array of social and economic indicators as towns move up the urban scale. It is however important to stress that the obvious linear relationship between urban size and other variables is contradicted by the smallest towns belonging to class VI (less than 5.000 inhabitants). For most variables, these micro-towns appear much more similar to cities than to small towns of say 5.000-20.000 inhabitants. The reason for this divergence is that these small towns are in most cases small localities adjacent to urban agglomerations which have been classified as urban by census authorities. Though very small in size, they share most of the social or economic characteristics of cities. These micro-towns typically belong to the urban fringe described by Hans Schenk in his paper.

2. MODELLING THE EFFECT OF SIZE ON URBANIZATION CHARACTERISTICS

Though our data are poor on decentralized industrialization as such, we have enough information to describe the characteristics of towns according to their demographic size as the previous tables and graphs have shown. We will go here

beyond the mere description of these variables to provide a more analytical approach of the impact size on urbanization. The first analysis concerns the statistical distribution of urban places by size.

2.1. Rank size curves

Figure 4 is based on the standard rank-size relationship. Towns are ranked by descending population size and both ranks and populations are plotted on a log scale. The idea behind these figures, also known as Zipf curves, is that urban populations often exhibit an almost regular pattern. The slope of the curve provides an indication of urban unbalanced structure. As Indian towns show on Figure 5, the distribution of towns by size is rather regular, with a slope very close to (minus) one.

There are however obvious irregularities at both the ends of the graph. On the left side of the curve, bigger cities (million-plus cities) do not fit the straight line as they display strong individual characteristics inherited from their own demographic and historical momentum. Moreover, such metropolises usually define sub-regions of their own and determine independent urban structures or "urban fields". For other reasons, there is an apparent deficit of smaller towns on the right side of the curve. This is partly due to the exclusion of rural settlements (villages) from the analysis. At the same time, the Census of India has included within its definition of urban places small towns of less than 10,000 inhabitants, including towns with a strong agricultural orientation or towns that belong to the rural fringe of metropolitan agglomerations as has been previously observed.

Similar graphs have been prepared to take care of these different problems by excluding urban areas with population above 1,000,000 or below 10.000 and by distinguishing between macro-regions (East, South, Northwest). These curves are reproduced on Figure 6.

2.2. Factor analysis

The large variety of available variables (more than 30 variables) suggests the use of a factor analysis aimed at data reduction. The idea behind this procedure is to highlight the most significant dimensions of the global urbanization process. The results of this tentative factor analysis (using the principal component method) are presented on Table 9. Here, we have used almost all variables available except town size. Only three factors have an eigenvalue greater than one. As the first factor proves to be extremely powerful (eigenvalue greater than 4) and account for almost a fourth of the total variance, we will restrict our attention to this factor. The values given on the second part of Table 6 are the most significant correlation coefficients, both positive and negative, between the factors and the original variables.

As can be seen, this factor brings together a bunch of variables that correspond to high literacy, quality of housing and almost all the economic

activities outside the primary sector. This factor is also negatively correlated to large families and agricultural occupations. To some extent, this factor sums up the attributes expected from a "modern" town and we will therefore label it the "modernization factor". Modernization is both a more complex phenomenon and a very vague label within the social sciences, but the term is retained here simply for want of a better one. Modernization as measured by our coefficients could also be expressed in terms of "urban quality", i.e. quality of human capital or of housing. This would however fail to reflect the economic component (shift from primary towards secondary and tertiary sectors).

On Figure 2, this factor is plotted for individual towns against their population. This graph takes its significance from the fact that there is a distinct, positive statistical relationship between urban size and "modernization". However, the micro-towns that have been already hinted at tend to blur this relationship as they possess attributes closer to those of cities than of small towns. As already observed, they are much more "modernized" than urban areas of similar size. The curve fitted reflects actually the relation between modernization and size for towns above 5,000 inhabitants. Towns located above the curve can be considered as privileged towns as their position is better in terms of literacy, urban amenities, etc. Inversely, towns plotted below the fitted curve are somewhat backward.

2.3. Modelling urban growth

As a last illustration of the data base's potentialities, we will examine the process of urban dynamics as summarized by the 1981-91 growth rate for individual towns[1]. Here, we are presenting the results of a basic step-wise least-square regression with urban growth rates as a dependent variable. No effort has been made to modify variables (via a log transformation, etc.) or to weigh individual data (by population, etc.). However unrefined the model may appear, the results shown on Table 10 are quite interesting as they reflect the expected components of growth as well as lesser known correlates.

Urban growth is the product of different phenomenona. The first relates directly to the administrative definition of individual towns as from census to census, town boundaries may increase because of the absorption of neighbouring townships or villages. The impact of urban redefinition is captured by a specific variable ("area growth", i.e. the spatial growth rate) and proves to be very significant. Similarly, one of the first demographic components (viz. fertility) is also closely related to urban growth. It turns out to be actually the most important single correlate of urban growth. There are however no way to find indirect estimates of mortality, but this factor is both less important in absolute terms and less

1. This rate is not available for smaller units that gained the urban status in 1991. There are other towns (usually of smaller size) for which the growth rates could not be computed.

heterogeneous between towns. It is therefore unlikely to be a major determinant of urban dynamics.

The remaining demographic component of urban growth is of course the intensity of migration, which gets marginally reflected in the overall sex ratio as labour migrations are often masculine. Naturally, migration also correlates strongly with the living and working conditions of the individual towns. Interestingly enough, the global modernization factor has been found to be insignificant in this model compared to individual variables and has therefore been dropped from the regression.

Among the variables most correlated with urban growth (and its migratory component), participation rates for both men and women confirm the role played by labour migration and the economic character of urban attractiveness. Several economic sectors are also shown to be closely associated with urban growth, most notably the other services sector and the industrial sector. The construction sector also appears to be a good indicator of urban dynamics although it is a rather small component of the urban workforce (see Table 6). The structural impact of town size (log population) on urban growth is present, though of limited strength.

Table 4. Social and Demographic Variables (Indian Towns 1991)

Variable	Mean	Std. Dev	Min Val	Max Val	Nb of towns 10% above the mean
Sch. Castes	0.135	0.084	0.000	0.952	366
Sch. Tribes	0.043	0.114	0.000	0.997	284
Sex ratio	910.754	81.692	0.000	1590	297
Fertility	0.440	0.106	0.000	0.853	724
Literacy					
Men	0.767	0.121	0.231	0.997	704
Women	0.549	0.167	0.059	0.997	1075
Total	0.664	0.139	0.166	0.997	878
Househ/ res	1.039	0.063	1.000	1.880	216
Pop/ househ	5.640	0.948	1.370	10.310	927
Pop/ resid	5.878	1.101	1.407	11.078	968
Density	4473	6389	78	95974	1010

Sources: *Census 1991*, PCA files
Sample: *3697 urban units* (urban agglomerations, towns)

Table 5. Female and Male Economic Activities

Variable	Mean	
	Women	Men
Cultivators	0.112	0.132
Ag.laborers	0.243	0.116
Livestock etc	0.017	0.026
Mining	0.012	0.019
Household Ind	0.081	0.037
Other Industries	0.084	0.135
Construction	0.020	0.043
Trade etc.	0.070	0.212
Transport etc	0.009	0.066
Other Services	0.353	0.213
Total Total	1.000	1.000
participation rate	0.110	0.575

Table 6. Economic Activities

Variable (both sexes)	Mean	Std. Dev	Min	Max	Nb of Towns 10% Above the Mean
Cultivators	0.131	0.126	0.000	0.900	596
Ag.labourers	0.139	0.117	0.000	0.725	721
Livestock etc	0.025	0.051	0.000	0.906	90
Mining	0.019	0.090	0.000	0.854	129
Household Ind	0.043	0.078	0.000	0.895	194
Other Industry	0.129	0.114	0.000	0.951	462
Construction	0.039	0.039	0.000	0.755	68
Trade etc.	0.193	0.088	0.000	0.734	425
Transport etc	0.058	0.047	0.000	0.709	84
Other Services	0.223	0.118	0.014	0.976	524

NB: No of towns with more 25% workers in agriculture: **1728**
No of towns with more 50% workers in agriculture: **526**

Table 7. Town Size, Amenities, Literacy and Scheduled Castes

Town class	Electricity	Toilets	Sch. Caste	Literacy
VI	.717	.504	0.155	0.733
V	.620	.413	0.143	0.626
IV	.645	.450	0.139	0.639
III	.677	.515	0.127	0.672
II	.754	.595	0.116	0.706
I	.809	.719	0.116	0.759
Mean	.767	.649	0.120	0.728

Table 8. Town Size and Occupation

Town class	Agriculture	Trade	Other services	Non-household Industry
VI	.255	0.156	0.280	0.109
V	.340	0.166	0.201	0.108
IV	.319	0.187	0.203	0.119
III	.255	0.208	0.222	0.131
II	.141	0.235	0.256	0.165
I	.049	0.228	0.269	0.256
Total	.116	0.221	0.254	0.214

Table 9. Principal Component Analysis

(principal components) Component	Eigenvalue	Difference	Proportion	Cumulative
1	4.20113	2.05035	0.2211	0.2211
2	2.15078	0.41595	0.1132	0.3343
3	1.73483	0.38691	0.0913	0.4256

Table 10. Description of Factors 1 and 2

Variable 1 (modernisation)		Variable 2	
Literacy	0.40168	Pop/household	0.45567
Electricity	0.34853	Fertility	0.39397
Toilets	0.29850	Trade	0.34158
Services	0.27233	Toilet	0.23100
Transport	0.24590	Services	0.18600
Construction	0.20015		
Trade	0.19844		
Other industries	0.16045		
Pop/household	-0.22898	Sc. tribe	-0.18124
Fertility	-0.28347	Litteracy	-0.19453
Ag. labourers	-0.30233	Ag. lab.	-0.22628
Cultivators	-0.35043	Fem/males	-0.31237
		Partic.rate	-0.40481

Table 11. Determinants of Urban Growth

procedure
Method: Step-wise regression:
Dependant variable: 1981-91 population growth rate
Independent variables: Social data, Urban size, Workforce composition etc.
Sample: 1285 Towns (outliers excluded)
stepw pgr cwr agr pop srt scaste stribe fliterat mlit fworker mworker
> tw5* tag tw4- tw9 if pg<2 & agr<2,forw

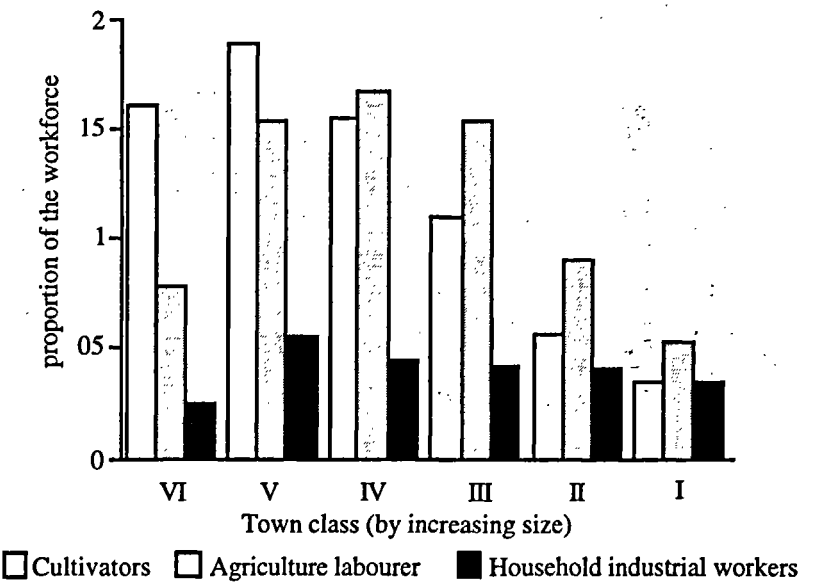
Model

Source	SS	df	MS
Model	8.0167755	16	.501048465
Residual	25.6460452	1268	.020225588
Total	33.6628206	1284	.026217150

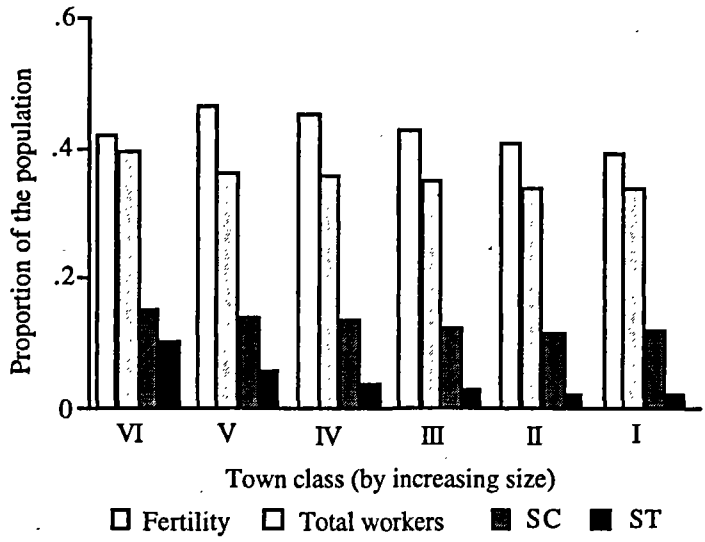
Observations = 1285
F(16, 1268) = 24.77
Prob > F = 0.0000
R^2=0.2381/Adjusted R^2 = 0.2285
Root MSE = 0.14222

Population growth	Coefficient	Std. Err	t	P>|t|
Fertility	.7851708	.0807969	9.718	0.000
Area growth	.0881351	.0137355	6.417	0.000
Population (log)	.0123423	.0053528	2.306	0.021
Sex ratio (fem/males)	-.0002342	.0000995	-2.353	0.019
Scheduled tribes	.1024129	.0634606	1.614	0.107
Female literacy	.1234267	.0527924	2.338	0.020
Female part. rate	.222698	.0850832	2.617	0.009
Male part. rate	.4724751	.1030965	4.583	0.000
Household industry	.1273362	.1113044	1.144	0.253
Other industries	.2337042	.1016958	2.298	0.022
Agriculture	.1045574	.0948163	1.103	0.270
Mining	.3442355	.1171525	2.938	0.003
Construction	.7350212	.1772457	4.147	0.000
Trade	.0850594	.1146615	0.742	0.458
Transport	.1815942	.1352537	1.343	0.180
Other services	.3124284	.1090031	2.866	0.004
Constant	.3959258	.1885200	2.100	0.036

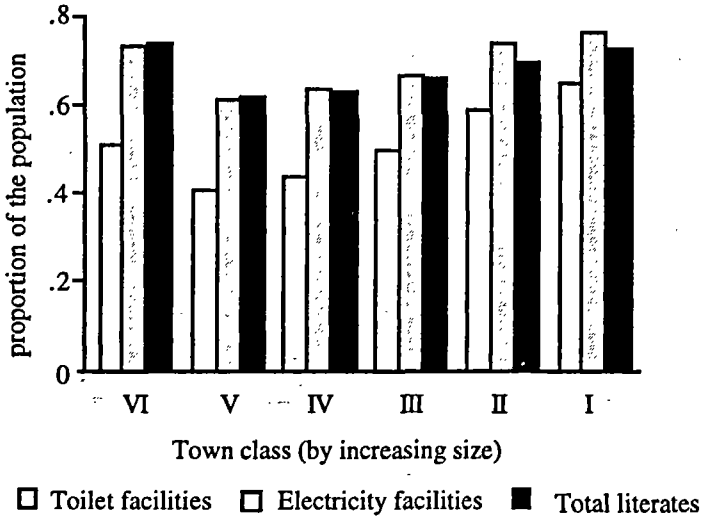
Graph 1. Proportion of the Workforce According to Town Size (Agriculture and Household Industry)



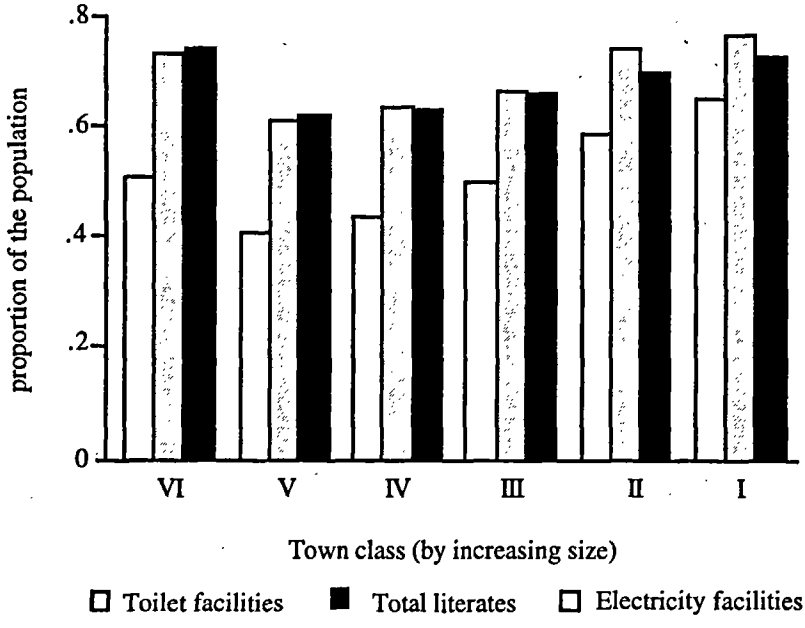
Graph 2. Proportion of the Population according to Town Size(Participating Rate, Scheduled Castes and Tribes) and Average Fertility Level (Child/Women Ratio)



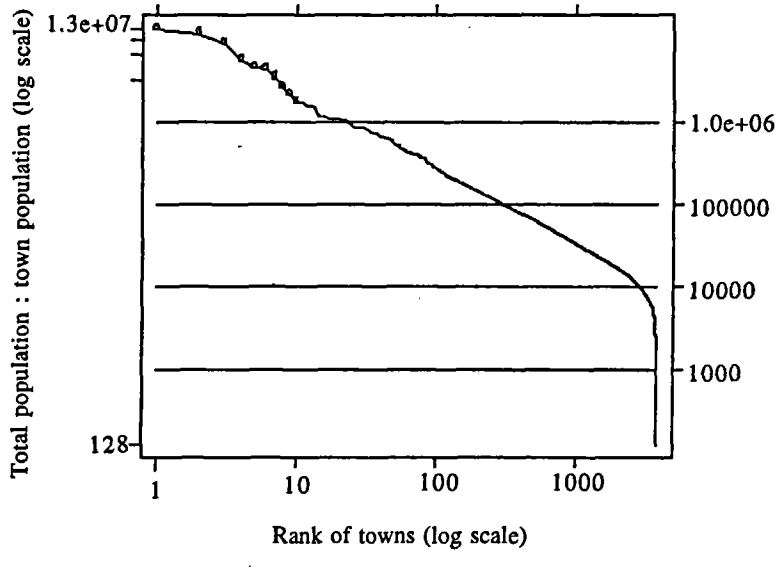
Graph 3. Proportion of the Workforce According to Town Size (Transport, Other Industries, Trade and Commerce, Other Services)



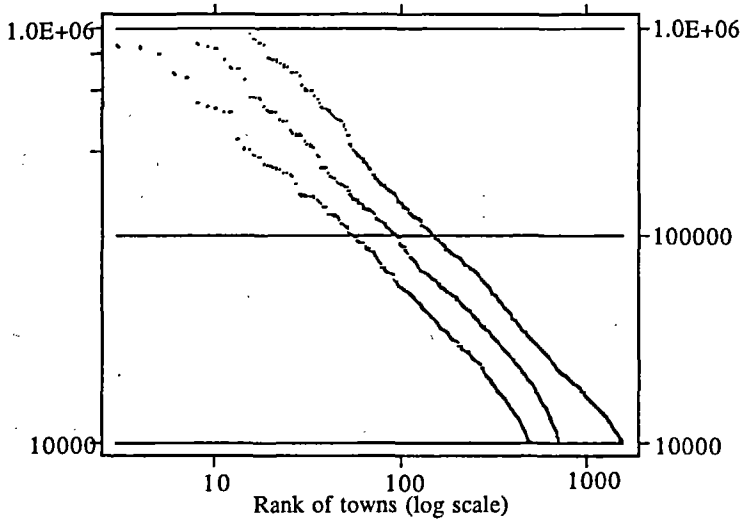
Graph 4. Proportion of the Population According to Town Size (Amenities and Literacy)



Graph 5. Populations of Towns ranked by decreasing Size (Zipf Curve)



Graph 6. Populations of Towns ranked by decreasing Size for East, South and North-West regions (Zipf Curves)



Graph 7. Modernization factor and population (individual towns and fitted curve)