

## Appendix I

### **Notes on the Data: Adjustment of the Data into Bhalla and Singh's District Units, and Estimation by Interpolation or Extrapolation**

In this study, as the basic geographical unit of aggregation of statistical data we use the district or districts combined as defined in Bhalla and Singh's study.<sup>1</sup> All the geographically defined statistical data are re-distributed and aggregated according to the district boundaries used in their study. Then, the geographically adjusted data are linearly interpolated or extrapolated in order to find estimated values for the election years (1957, 1962, 1967, 1971, 1977, 1980, 1984, 1989, 1991, 1996, 1998, and 1999).

However, before beginning the adjustment and linear estimation, there are many problems to deal with concerning the data sets themselves, such as the inconsistency of definitions between Censuses, etc. In this Appendix, we explain the problems with the data sources and the ways we dealt with these problems. Following that, some notes concerning the data adjustment and estimation are shown.

#### **1. Bhalla and Singh's Agricultural Data**

Bhalla and Singh's agricultural data are triennium averages for the periods of 1962–65, 1970–73, 1980–83, and 1990–93. Averaging the values for three years is expected to eliminate most of the yearly fluctuations from the data. The data, therefore, are considered to show the middle-term trend concerning the agricultural production.

The data are converted by linear interpolation or extrapolation, into data sets for the years of Lok Sabha elections. In the process of estimation, some of the estimated values for 1957 and 1962 take minus values as a result of the extrapolation (For example, "Fertiliser Consumption"), which is very unnatural. In these cases, values of zero are substituted.

## 2. Population Census Data

The most serious problem in the comparative analysis of the Census data of different points of times is the change of the concepts and definitions between Censuses. These problems are examined briefly in each case.

### (1) Literacy

A person is identified as literate if he or she can both read and write in at least one language. This definition of “literacy” is basically maintained from the 1961 to 2001 Censuses, on which this study is based. However, the literacy rates in 1961, 1971 and 1981 relate to the population aged five years and above, while those for the 1991 and 2001 Censuses relate to population aged seven years and above.<sup>2</sup> The literacy rates in the 1991 and 2001 Censuses might have been slightly lower if the same age definition as in the 1961, 1971 and 1981 Censuses had been adopted. Although the potential difference in literacy rate owing to the change of age limit may not be a large one, it is better to adopt a more comprehensive definition in order to alleviate the definitional problem. In this study, the “crude literacy rate” is adopted for this purpose, but the problem of the definitional changes cannot be completely avoided. Literacy is defined as follows:

Crude literacy rate (percentage)

$$= \frac{\text{Number of literates irrespective of ages (exclusive of 0-4 or 0-6 age group)}}{\text{Total population}} \times 100$$

The consistency of the “crude literacy rates” between 1961, 1971 and 1981 Censuses on the one hand, and 1991 and 2001 Censuses on the other is higher than that of the “literacy rate” between the two Census groups. This is because the denominators are the “Total population” and there are few literate children in the 5-6 age group. The crude literacy rates are calculated for both males and females.

### (2) Urbanisation

According to the 2001 Census, urban areas are defined as follows:

- a) All statutory areas with a municipal corporation, cantonment board, municipality, notified town area committee, nagar panchayat etc.

- b) All other areas which satisfy the following criteria :
- i) A minimum population of 5,000
  - ii) At least 75 percent of the male working population are engaged in non-agricultural pursuits and
  - iii) A density of population of at least 400 per Sq. Km.

This definition of urban area has remained basically unchanged since the 1961 Census. However, there have been some small changes. The most relevant concerned the content of the concept of ii) of b) above. Concerning the “non-agricultural pursuits,” the occupations of forestry, fishing, livestock, hunting, logging, plantations and orchards etc., were excluded in the 1981, 1991 and 2001 Censuses. However, the impact of this marginal change in the definition is considered to be negligible.<sup>3</sup> The definitional substance of “Urbanisation,” therefore, can be considered to be basically continuous from the 1961 to 2001 Censuses. “Urbanisation” is defined as the percentage of the population in urban areas vis-à-vis the total population.<sup>4</sup>

### **(3) *Scheduled Castes (=SCs) and Scheduled Tribes (=STs)***

SCs and STs are identified by the Union Parliament on the basis of consultations with the State government concerned.<sup>5</sup> The State government, therefore, actually plays an important role in the process of inclusion or exclusion of the name of a caste or tribe from the list of the schedule. Therefore, a certain caste or tribe can be a SC or ST in one State but not in another. Due to this, there can be some inconsistency in the coverage of the concepts between States. In addition, certain castes or tribes has sometimes been added to or removed from the list of SCs or STs. Because of this, there can be inconsistencies in strict comparisons of the percentage of SCs and STs between Censuses even in the same State.

However, the inconsistency of the concepts between States, and the inconsistency owing to the change in the composition of the list of SCs or STs between Censuses, are not considered to be essentially damaging since the basic criteria of the identifications is not altered. The criteria are social disability and discrimination on the basis of untouchability in the past and socio-economic backwardness in the case of SCs, and of cultural peculiarity, socio-economic backwardness, and geographical isolation in the case of the STs. In this study, we adopt the percentage of SCs or STs vis-à-vis the total population without any adjustment concerning the two inconsistencies mentioned above. However, it is better to take these inconsistencies into account while interpreting the statistical outcomes.

The SCs or STs population data from the 2001 Census were still not available for the author in April, 2003.

The following two tables are the basic statistics on the percentages of SCs and STs, on the basis of the District boundaries in Bhalla and Singh's study. They are calculated in order to make a rough check on the continuity of the Censuses. Both tables show the very high level of continuity of the data set.

**Table AI-1**  
**Basic Statistics on the "Percentage of SCs"**

Census Year	mean (%)	Standard deviation (%)	N	Pearson' correlation coefficient	
1961	15.1	7.3	273		
1971	15.2	7.3	273	1961 and 1971	0.987
1981	16.1	7.2	273	1971 and 1981	0.967
1991	17.0	7.2	273	1981 and 1991	0.917

Note: 1) The samples are calculated in line with the District boundaries in Bhalla and Singh's study.

2) Samples with missing values are excluded.

3) All the correlation coefficients are statistically significant at the level of 0.01.

**Table AI-2**  
**Basic Statistics on the "Percentage of STs"**

Census Year	mean (%)	Standard deviation (%)	N	Pearson' correlation coefficient	
1961	8.6	16.3	273		
1971	8.7	16.2	273	1961 and 1971	0.996
1981	10.0	16.1	273	1971 and 1981	0.976
1991	10.0	16.1	273	1981 and 1991	0.971

Note: Same as Table AI-1.

#### **(4) Non-Hindu Population**

The non-Hindu population is calculated by subtracting the percentage of the Hindu population from 100 percent. The religion-wise population data from the 2001 Census was still not available in April 2003. The linear estimation of values in the election year is easy because the percentage of the religion-wise population has remained basically stable since the 1961 Census.

### (5) *Cultivators and Agricultural Labourers*

The definitional changes of cultivators and agricultural labourers are problematic in the population data. However, the basic conceptual definition is, by and large, consistent. A person is a "cultivator" if he/she is engaged in cultivation independently or under supervision or direction as the owner or lessee of land held from the Government or as a tenant on land held by a private person or institution for payment of money, kind or share. A person is an "agricultural labourer" if he/she works on another person's land for wages in money, kind or share, and has no right of lease or contract on the land on which he/she works. However, the way to identify them has frequently changed. In this study, the two variables are calculated as follows;

#### Percentage of Cultivators

$$= \frac{\text{Number of Cultivators}}{\text{Number of Workers}} \times 100 : \text{In the 1981, 1991, and 2001 Censuses}$$

$$= \frac{\text{Number of Cultivators}}{\text{Number of Main Workers}} \times 100 : \text{In the 1961 and 1971 Censuses}$$

#### Percentage of Agricultural Labourers

$$= \frac{\text{Number of Agricultural Labourers}}{\text{Number of Workers}} \times 100 : \text{In the 1981, 1991, and 2001 Censuses}$$

$$= \frac{\text{Number of Agricultural Labourers}}{\text{Number of Main Workers}} \times 100 : \text{In the 1961 and 1971 Censuses}$$

In the 1961 and 1971 Censuses, the total population is dichotomised into "workers" and "non-workers." A "worker" is a person whose main activity is participation in any economically productive work involving physical or mental activity. Work involves not only actual work but also the effective

supervision and direction of work. Cultivators and Agricultural Labourers are subdivisions of “workers.” However, the methods of identification of “workers” and their economic activities differ slightly between the two Censuses.<sup>6</sup>

In the 1981 and 1991 Censuses, the total population is dichotomised into “workers” and “non-workers,” and the former category further dichotomised into “main workers” and “marginal workers.” “Main workers” are subdivided into many economic groups such as Cultivators, Agricultural Labourers, etc. There is no major subdivision in the case of “marginal workers.” In the 2001 Census, the total population is dichotomised into “total workers” and “non-workers,” and the former category is further dichotomised into “main workers” and “marginal workers.” The “main workers” in 1981, 1991, and 2001 are subdivided into many economic groups such as Cultivators and Agricultural Labourers. There are no subdivisions for “marginal workers” in these Censuses.<sup>7</sup> The simplified definitional changes of workers and others between Censuses are shown in the following table.

**Table AI-3**  
**Definitional Changes of “Workers”**

Census Year	Reference period	Period of Work Required (PWR) in reference period for Regular/Main Worker	PWR in reference period for Marginal Worker	PWR in reference period for Seasonal Worker
1	2	3	4	5
1961	For regular worker 15 days preceding the day of enumeration and for seasonal worker, working season.	At least one day in 15 days prior to enumeration.	-	An hour's regular work a day throughout the greater part of working season.
1971	For regular worker, 7 days preceding the day of enumeration and for seasonal worker one year preceding the day of enumeration.	At least one day in one week prior to enumeration.	-	-
1981	-do-	6 months (183 days) or more	Less than 6 months (183 days)	-
1991	-do-	-do-	-do-	-
2001	-do-	6 months or more	Less than 6 months	-

Source: Bhaskar, V.S. (Director of Census Operations, Andhra Pradesh). 2002. *Census of India 2001, Series-29, Andhra Pradesh, Provisional Population Totals, Paper-3 of 2001: Distribution of Workers and Non-Workers*. Delhi: Controller of Publications, p. 20.

Note: “Seasonal workers” includes, for example, those engaged in cultivation, livestock, dairying, household industry, etc.

The most problematic definitional change is considered to be the division of the “workers” category into “main workers” and “marginal workers.” Concerning the comparability of definition of “workers” between Censuses, the 1991 Census explained:

“As regards comparability of economic data of 1981 and 1991 censuses with those of 1971 and 1961, it is expected that the main workers would approximately correspond to the workers of 1971 while the main workers and marginal workers together would correspond to the workers of 1961.”<sup>8</sup>

The definition of “main workers” in the 2001 Census is basically consistent with that of the 1991 Census. It can, therefore, be said that the percentages of Cultivators and Agricultural Labourers calculated in the above formula can be expected, by and large, to be basically continuous between Censuses.

The following two tables present the basic statistics of the percentages of Cultivators and Agricultural Labourers on the basis of the District boundaries from Bhalla and Singh’s study, which are worked out in order to make rough check on the continuity of the Censuses. Both tables show a very high level of continuity of the data set.

**Table AI-4**  
**Basic Statistics on the “Percentage of Cultivators”**

Census Year	mean (%)	Standard deviation (%)	N	Pearson’ correlation coefficient	
1961	55.6	15.7	271		
1971	47.5	15.4	271	1961 and 1971	0.955
1981	46.2	16.0	271	1971 and 1981	0.942
1991	42.8	16.1	271	1981 and 1991	0.929
2001	35.2	15.2	271	1991 and 2001	0.894

Note: 1) The samples are calculated in line with the District boundaries in Bhalla and Singh’s study.

2) Samples with missing values are excluded.

3) All the correlation coefficients are statistically significant at the level of 0.01.

**Table AI-5**  
**Basic Statistics of the “Percentage of Agricultural Labourers”**

Census Year	mean (%)	Standard deviation (%)	N	Pearson’ correlation coefficient	
1961	15.8	10.0	271		
1971	25.4	11.6	271	1961 and 1971	0.933
1981	23.7	11.4	271	1971 and 1981	0.894
1991	25.2	11.3	271	1981 and 1991	0.868
2001	26.9	12.2	271	1991 and 2001	0.806

Note: Same as Table AI-4.

### **(6) Adjustment and Estimation**

The data from the population Censuses in 1961, 1971, 1981, 1991 and 2001 are adjusted in line with the boundaries of Bhalla and Singh's district units. The procedure is explained in Section 2 of Chapter 2. After the allotment, the demographic data for 1961, 1971, 1981, 1991 and 2001 are converted by interpolation or extrapolation into the data sets in the Lok Sabha election years. In the process, missing values are also estimated if possible. The data for the Assam in the 1981 Census, for example, are missing because the Census enumeration could not be conducted due to the ethnic turmoil. The data for Assam in 1981 are estimated from the data in 1971 and 1991.

For the ST ratio and other variables, several values are estimated to be below zero in 1957, 1998, or 1999, as a result of the extrapolation by linear estimation. In these cases, the value of zero is substituted.

### **3. Lok Sabha Election Data**

Major delimitation exercises were conducted on three occasions after the State reorganisation in 1956. Each Lok Sabha constituency consists of several State Legislative Assembly constituencies, as has been explained already. After 1956, therefore, Lok Sabha and State Legislative Assembly elections are grouped into elections in the periods between 1956 to 1966, 1966 to 1976 and from 1976 onwards in terms of the boundaries of both Lok Sabha and State Legislative Assembly constituencies. Three conversion formulas are necessary to convert the Lok Sabha election data sets to Bhalla and Singh's district boundary based data sets.

Assembly Constituencies are the smallest geographical units whose combination forms, in most cases, a district or Parliament Constituency. The materials on which these relationships are identified are; Election Commission of India 1957, Delimitation of Parliamentary and Assembly Constituencies Order, 1956, New Delhi, 1957; —, 1967, Delimitation of Parliamentary and Assembly Constituencies Order, 1966, New Delhi, 1967; —, 1976 Delimitation of Parliamentary and Assembly Constituencies Order, 1976, Government of India Press, New Delhi, 1976.

The first step is to estimate the number of electorates of all the State Legislative Assembly constituencies in Lok Sabha election years by linear interpolation or extrapolation. The years for the original State Legislative Assembly election data are shown in the three tables below. For the periods from 1956 to 1966, the 1957 data of the electorate of the State Legislative

Assembly is used with the exception of the Andhra region of Andhra Pradesh. The State Legislative Assembly election was conducted in 1962 in most of the States. However, the boundaries of the State Legislative Assembly constituencies in the 1962 election are not completely the same as that of the 1957 election because of the bifurcation of the two-member constituencies and the minor boundary changes accompanying the bifurcation, which were made just before the 1962 election. The changes in the boundaries of the State Legislative Assembly constituencies were especially extensive in the Gujarat, which was created as a result of the bifurcation of Bombay State in 1960. The data on the electorates of the State Legislative Assembly constituencies in 1957 are utilised as weight for the 1957 and 1962 Lok Sabha elections, with the exception of the Andhra region of Andhra Pradesh. In the latter case, the State Legislative Assembly election was not conducted in 1957 but in 1955, and the changes of boundaries caused by the bifurcation of the two-member constituencies were minimum. The 1955 and the 1962 data, therefore, were utilised for the Andhra region. In the other two periods, the electorate data from the two State Legislative Assembly elections were adopted in each State, as shown in the tables.

The second step is the conversion of the Lok Sabha constituency based data set into the data sets based on Bhalla and Singh's districts. The Lok Sabha election data are allotted to Bhalla and Singh's district units on the basis of the weights of the number of the electorate in the State Legislative Assembly constituencies estimated in the first step. Thus, for the 1957 and 1962 Lok Sabha elections, the data based on 393 and 481 Lok Sabha constituencies each are converted into 279 data sets based on Bhalla and Singh's districts, with the weights of 2,420 State Legislative Assembly electorate data. In the cases of the 1967 and 1971 Lok Sabha elections, data based on 489 Lok Sabha constituencies are converted into 279 data sets based on Bhalla and Singh's districts, with weights of 3,168 State Legislative Assembly electorate data. For the 1977 Lok Sabha elections and onwards, data based on 507 Lok Sabha constituencies are converted into 279 data sets based on Bhalla and Singh's districts, with weights of 3,318 State Legislative Assembly electorate data. The details of the conversion have been already explained in the text.

It must be noted that minor boundary changes have been frequent. These include, for example, changes in the boundaries of towns, tehsils, development blocks, etc., which border on a Lok Sabha constituency and thus affect its boundaries. When a river separating two Lok Sabha constituencies changes course, for example, most probably the boundaries of the Lok Sabha constituencies will be modified accordingly. However, these

**Table AI-6**  
**Year of the Original State Legislative Assembly Election Data Based on the 1956**  
**Delimitation, Adopted to Get Estimations of the Electorate (=weight) in 1957 and 1962**

State	Year of the Original Legislative Assembly Election Data		
Andhra Pradesh	{ Andra Region	1955	1962
	{ Telengana	1957	-
Assam		1957	-
Bihar		1957	-
Bombay		1957	-
Kerala		1957	-
Madhya Pradesh		1957	-
Madras		1957	-
Mysore		1957	-
Orissa		1957	-
Punjab		1957	-
Rajasthan		1957	-
Uttar Pradesh		1957	-
West Bengal		1957	-

Source: State Legislative Assemblies Data before 1985 are from: Singh, V. B. and Shankar Bose. 1987. *States Elections in India: Data Handbook on Vidhan Sabha Elections 1952-85, Vol. I, II, III, IV, and V*. New Delhi: Sage; Data after 1985 are collected from: <http://www.eci.gov.in/ARCHIVE> (accessed: 9 October 2002).

**Table AI-7**  
**Year of the Original State Legislative Assembly Election Data Based on the 1966**  
**Delimitation, Adopted to Get Estimations of the Electorate (=weight) in 1967 and 1971**

State	Year of the Original Legislative Assembly Election Data	
Andhra Pradesh	1967	1972
Assam	1967	1972
Bihar	1967	1972
Gujarat	1967	1972
Haryana	1967	1972
Kerala	1967	1970
Madhya Pradesh	1967	1972
Madras	1967	1971
Maharashtra	1967	1972
Mysore	1967	1972
Orissa	1967	1971
Punjab	1967	1972
Rajasthan	1967	1972
Uttar Pradesh	1967	1969
West Bengal	1967	1971

Source: Same as above.

**Table AI-8**  
**Year of the Original State Legislative Assembly Election Data Based on the 1976**  
**Delimitation, Adopted to Get Estimations of the Electorate (=weight)**  
**in 1977, 1980, 1984, 1989, 1991, 1996, 1998, and 1999**

State	Year of the Original Legislative Assembly Election Data	
Andhra Pradesh	1978	1999
Assam	1978	2001
Bihar	1977	2000
Gujarat	1980	1998
Haryana	1977	2000
Karnataka	1978	1999
Kerala	1977	2001
Madhya Pradesh	1977	1998
Maharashtra	1978	1999
Orissa	1978	2000
Punjab	1977	2000
Rajasthan	1977	1998
Tamil Nadu	1977	2001
Uttar Pradesh	1977	1996
West Bengal	1977	2001

Source: State Legislative Assemblies Data before 1985 are from: Singh, V. B. and Shankar Bose. 1987. *States Elections in India: Data Handbook on Vidhan Sabha Elections 1952-85, Vol. I, II, III, IV, and V*. New Delhi: Sage; Data after 1985 are collected from: <http://www.eci.gov.in/ARCHIVE> (accessed: 9 October 2002).

minor changes do not affect the process of statistical analysis in practice because they only affect at most a small geographical area of the constituency. The more important problem for the conversion is the problem of aggregation. In the case of the 1977 Lok Sabha elections and onwards, for example, one Bhalla and Singh district unit consists of 1.82 (=507/279) Lok Sabha constituencies on average. In this process of aggregation, the particularity of each constituency is, in a sense, attenuated. Whether the attenuation is damaging to the study depends on the character of the data. For example, the voter turnout percentage is less affected than the percentage votes poll by parties, as shown in the text.

## Notes

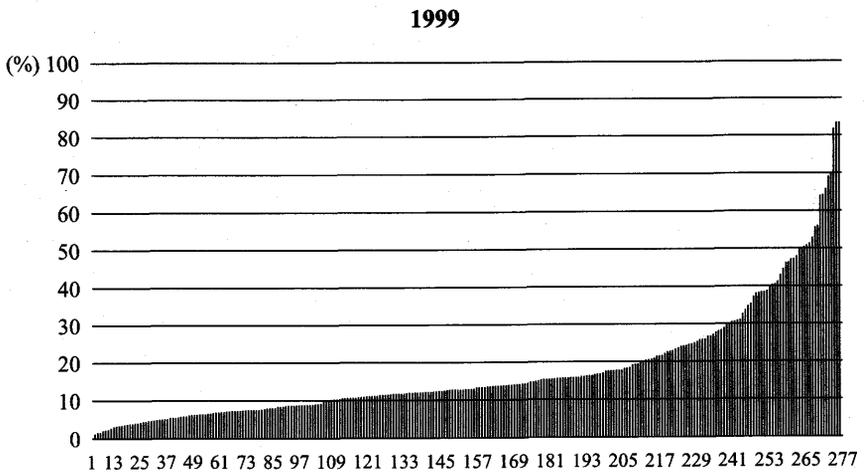
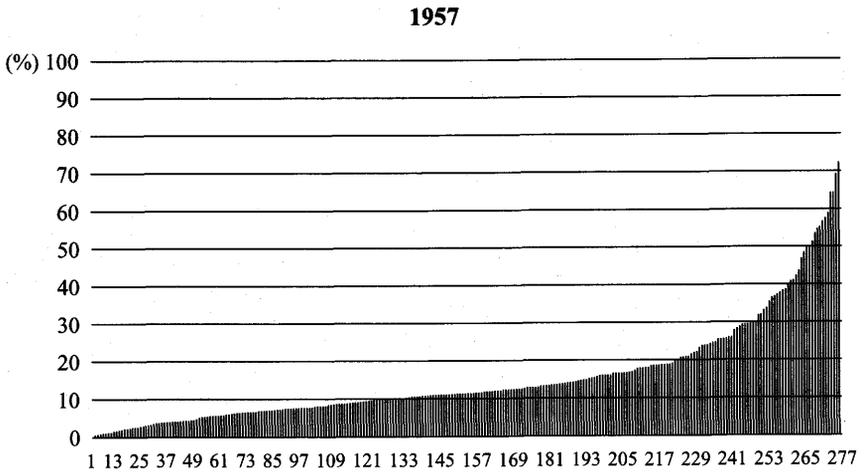
- <sup>1</sup> See, Bhalla, G. S. and Gurmail Singh. 2001. *Indian Agriculture—Four Decades of Development*. New Delhi: Sage, Appendix 1.1.
- <sup>2</sup> See, Nanda, A.R. (Registrar General & Census Commissioner, India). 1994. *Census of India 1991, Series 1, India, Part II B(i), Primary Census Abstract, General Population, Vol. 1*. Faridabad: Government of India Photolitho Press, p. xlvii.
- <sup>3</sup> Office of the Registrar General (Ministry of Home Affairs, Government of India). 1986. *Census of India 1981, Census Monograph No. 1, Urban Growth in India 1951–81 (A Statistical Analysis)*. Delhi: The Controller of Publication, p. 1.
- <sup>4</sup> The introduction of one new concept may have affected slightly the definition of the concept, “urbanisation”. The concept of “Urban Agglomeration (U.A.)” was adopted in the 1971 Census, replacing that of town-group adopted in the 1961 Census. It has been used up to the 2001 Census. U.A. includes; “(a) a city with a continuous out-growth (the part of outgrowth being outside statutory limits but falling within the boundaries of the adjoining village or villages); (b) one town with a similar out-growth or two or more adjoining towns with their out-growths as in (a); and (c) a city and one or more adjoining towns with their outgrowths all of which form a continuous spread.” The concept includes parts of “the adjoining village or villages,” which can be a marginal extension of the concept of urban area. However, this marginal change of definition is not serious. See, Sekhar, A. Chandra (Registrar General & Census Commissioner, India). 1975. *Census of India 1971, Series I, India, PART II-A (i), General Population Tables*. Delhi: The Controller of Publications, p. 4.
- <sup>5</sup> See, *The Constitution of India*, Article 341 and 342.
- <sup>6</sup> See, Sekhar, A. Chandra (Registrar General & Census Commissioner, India). 1977. *Census of India 1971, Series I, India, PART II-B (i), General Economic Tables (B-I Part A & B-II)*. Delhi: The Controller of Publications, p. vi
- <sup>7</sup> Concerning the changes in these categorisations, see for example, Bhaskar, V.S. (Director of Census Operations, Andhra Pradesh). 2002. *Census of India 2001, Series-29, Andhra Pradesh, Provisional Population Totals, Paper-3 of 2001: Distribution of Workers and Non-Workers*. Delhi: Controller of Publications, p. 21.
- <sup>8</sup> See, Nanda, A.R. (Registrar General & Census Commissioner, India), *op. cit.*, p. lix.

## Appendix II

Figure AII-1

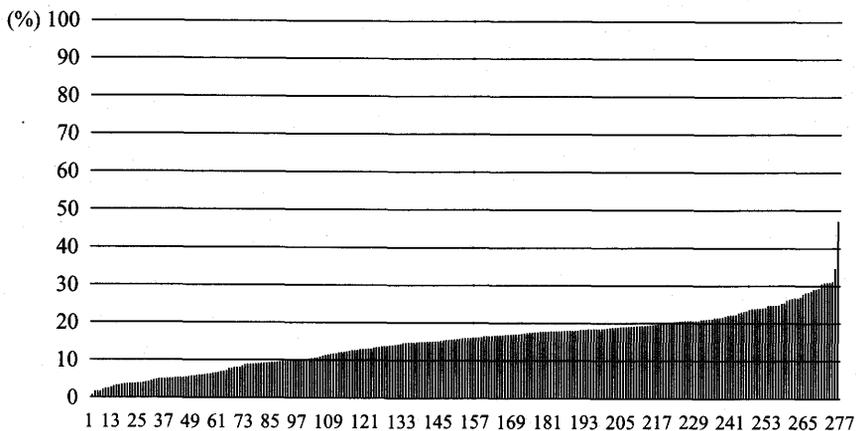
### Distributions of the Non-Hindu Population, the SC and ST Population

Distributions of the Non-Hindu Population  
Based on Bhalla and Singh's Districts Based Data Set

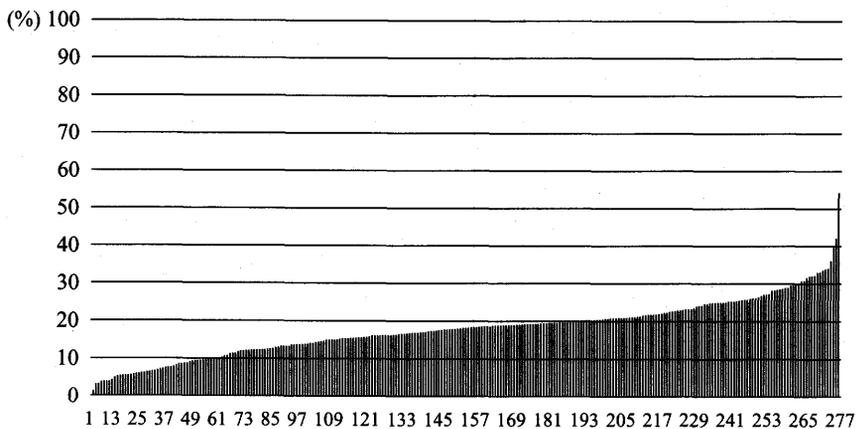


**Distributions of the SC Population  
Based on Bhalla and Singh's Districts Based Data Set**

**1957**

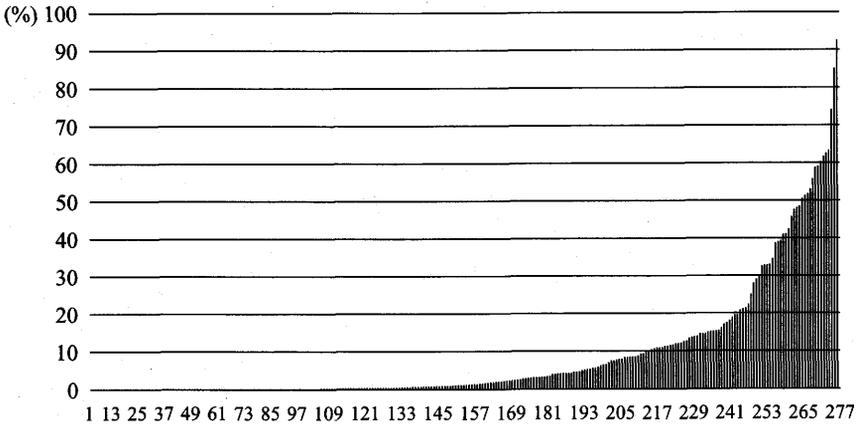


**1999**



**Distributions of the ST Population  
Based on Bhalla and Singh's Districts Based Data Set**

**1957**



**1999**

