Appendix I

THE DATA BASE

Reliability of the empirical result depends on the nature, accuracy and consistency of the data used for the analysis. The data requirements of the present study consist of data for analysing (i) industrial growth and trends thereof across states between 1961 and 1978, and (ii) evidence relating to various aspects of inter-state (inter-regional) variations in the level of industrial development in overall macro-economic terms. We outline the sources and nature of the data available for our purpose, in what follows.

1. The Data on Manufacturing Sector

The major source of data for the registered factory sector is Annual Survey of Industries (ASI), published annually by the Central Statistical Organisation (CSO), Government of India. The survey covers only registered segment of the manufacturing sector, under sections 2m(i) and 2m(ii) of the Factory Act 1948. For the purpose of the survey, factories in ASI frame are divided into two

1/ It may be mentioned that we have analysed only registered factory sector or registered manufacturing sector in the inter-regional context of India.
sub-sectors viz. census and non-census (sample). Generally, factories employing 50 or more workers and using power or employing 100 or more workers without using power constitute census sector and are completely enumerated in the survey, while registered factories employing 10 to 49 workers and using power or 20 to 99 workers without using power are covered on the basis of 50 per cent probability sample annually. The information about large scale factories (census sector) is available on a time series basis for the years between 1961 and 1978-79, except for the year 1972 when survey was not conducted. No separate information about sample sector is available after 1971-72. Moreover, the data for the period 1973-74 to 1978-79 are provisional ones.

The information about unregistered or unorganised segment of the manufacturing sector is extremely scarce. From no source, data on unorganised sector are available on regular or comparable basis (Somayajulu, 1975). Hence, analysis by and large, leaves out the unregistered or unorganised sector. However, we have used some information, provided by the National Accounts statistics, for this sector.

However, now information is available for years now, i.e. for 1979-80 to 1981-82. At the time of our actual analysis, latest data were available upto 1978-79 only.
1.1 The Limitations of the Data on Registered Factory Sector: The limitations in terms of availability and comparability over time as well as across industries have been extensively documented in the literature. Some of the major data limitations are:

(i) The annual survey of industries up to 1965 was conducted for the calendar year while afterwards, it is being conducted for the financial year.

(ii) All the factories covered under the Factory Act do not submit returns punctually. Consequently, some proportion of the registered factory sector remains unreported every year. Since no correction is made from year to year for non-reporting units, the comparability over time becomes restricted.

(iii) If any industry group consists of less than three reporting units in a given year, it is included under the classification of industries not elsewhere classified; while the states having less than 3 reporting units in a particular industry group are pooled together. However, in the footnotes of the industry information, the number of reporting units in respective states is mentioned against their names.

1/ For example, see Krishna (1972); Srinivasan and Vaidyanathan (1972); CSO (1977 and 1979); Ahlawalia (1985) among others.
Another problem was posed in ASI data by change in the industrial classification from ASI to National Industrial Classification (NIC), which is used since 1973-74. As a result industries under NIC do not bear one to one correspondence with ASI in some cases.

The data provided by ASI are at current prices. Therefore comparison over time is not ideal without adjusting the series for price variations.

These limitations make comparison of industries over time and across states somewhat difficult. However, despite the limitations, the ASI data are the best available set for what is termed as registered manufacturing sector in the National Accounts Statistics. Therefore, ASI data have been employed for the purpose of our analysis, after making certain adjustments as follows.

1.2 Adjustment in the Industrial Data. In order to arrive at registered factory sector, we aggregated the census and the sample sector data for the years 1961 to 1971-72 (henceforth referred to as 1971). Since 1973-74, CSO has been publishing information on the registered factory sector as such.

For the purpose of aggregating pooled information, particularly at the states' level, we divided the aggregate
information, as given under the column "others", using proportion of value added of a state among the clubbed states and the number of units pertaining to the state in question as weights; where

\[ X_{ij} = \text{value of a particular variable } X \text{ (like fixed capital, employment or value added etc.) in } i^{th} \text{ industry group in the clubbed } j^{th} \text{ state;} \]

\[ VA_j = \text{aggregate value added pertaining to the state } j; \]

\[ n_j = \text{total number of factories mentioned against the state } j \text{ (in the foot note); and} \]

\[ V_j = \text{weight of the } j^{th} \text{ state among all the clubbed states.} \]

Then:

\[ V_j = \frac{VA_j \cdot n_j}{\sum_{j=1}^{n} VA_j \cdot n_j} \quad (1) \]

\[ X_{ij} = V_j \cdot X_i \quad (2) \]

It may be mentioned that in adopting this methodology there is a chance of over or under estimating the region specific values of a variable. However, to provide an overall check on the accuracy of the data, the figures relating to a specific industrial characteristic were
aggregated across all the industry groups and compared with the total of that characteristic pertaining to a particular state.

Since the Census Sector and the Sample Sector information can be matched only at the three digit level of disaggregation, the census sector data was aggregated at three digit level of ASI using a key of correspondence in the sample reports. Therefore, NIC sectors also were suitably aggregated to correspond with the earlier period. Though, as mentioned earlier, a one to one correspondence could not be made possible even at three digit level; it was matched to be best possible extent. In all, 53 sector classification of industry groups was arrived at for the purpose of the analysis. Similarly these sectors were further aggregated in order to arrive at use base and input base sectoral classification.

As the data on industries are available at current prices, we feel that the nominal values are somewhat unreliable as indications, given the persistant inflation. It is quite likely that the inflation will effect different industry groups differently. Therefore, any comparison

1/ The aggregation scheme is presented in Appendix 2.
2/ See Section B and C of the Appendix 2.
over time will be meaningless, if made on the basis of current prices. The analysis is, thus, carried out taking the series of fixed capital and value added in constant terms, with 1970-71 as the base. The series of value-added were deflated by the appropriate wholesale price indices as provided in Chandak (1970). There are known limitations in using such price indices, and the ideal way of deflating the series is to use double deflation method. For want of information, we have not been able to do so, and resorted to the single deflation method. Similarly, fixed capital series was deflated by the Index of plant, machinery and transport equipment for the respective years. Fixed capital also represents depreciated value of some other fixed assets like land and building etc. along with the other components. Therefore, using the aforesaid deflator may not be very desirable. However, since plant and machinery and transport equipments account for the major proportion of the fixed capital, we can use this price deflator as the nearest proxy.

1/ For example in adopting this method we implicitly assume that the prices of inputs change at the same rate as that of the output. As an illustration, the use of the single deflation method offsets the trends in the petroleum products group in particular since the output prices diverge markedly from the input prices (Ahuwalia, 1985: 178).

2/ It may be mentioned that no adjustments have been made to correct for the vintage involved in the plant and machinery.
At this juncture one more point may be made clear. The ideal way of deflating various region/state specific fixed capital or value added series is to use region specific, industry-specific, price deflators. But the dearth of such indices is too well known. Therefore, even for the states, we have used all India price deflators. Consequently, the estimates suffer to the extent the regional variations in the price movements across industry groups are not captured. Nevertheless, we are, at least, able to capture the inter-industry price movements, and are able to correct for the growth of the various states, emanating from the sectoral composition. However, one has to keep these limitations in mind while viewing at the results.

As mentioned earlier, the focus of the study is the organised or registered manufacturing sector. However, the data on unorganised sector has been used to understand the overall variations in the manufacturing sector and its sub-sectors viz. organised and unorganised; and relationship between these segments across states in the chapter 4. The data are obtained from the Estimates of State Domestic Product (ESDP):1960-61 to 1982-83, published by the CSO (1984); and the various issues of the National Accounts Statistics (NAS).
2. Other Data

In analysing the determinants of inter-regional industrial variations and testing the various hypotheses, we required variety of data on a number of variables, which offer some explanations for the inter-regional industrial variations and declining regional industrial inequalities.

Apart from our major data sources viz. ASI, NAS, and ESDP; we obtained information from the various government reports and documents. Thus, for working out inter-industry linkages (see Chapter 6) we used Input-Output table obtained from A technical Note on the Sixth Plan of India (1980-85), Planning Commission (1981). Similarly in order to compute state specific sectoral income elasticities, we obtained data from the 28th round of the National Sample Survey (NSS); and the Task Force Report published by the Planning Commission (1979).

In order to gauge the impact of some of the behavioural variables on the inter-regional variations, we selected some proxies to represent these variables. Thus, data related to the various sets of population was obtained from the various volumes of the Census of India; road statistics from the various issues of the Basic Road Statistics of India, Ministry of Shipping and Transport,
Government of India; information on length of railways in the various states was obtained from the various issues of *Indian Railway Year Book*, Ministry of Railways, Railway Board; on banking and credit deposit ratios from *Chart Book on Financial and Economic Indicators*, Reserve Bank of India (RBI) (1978); on forest, mining, agriculture, power, nandayaloo, trade unions, geographical area of the various states from the various issues of the *Statistical Abstract: India* (CSO); inter-urban distances from the *Road Map of India*, Geological Survey of India, (1981); on the performance of non-banking financial institutions from the *Report of Development Banking in India*, Industrial Development Bank of India (1980-81), *Operational Statistics* of 1981, various reports on *Currency and Finance*, RBI; on registration of companies with the Directorate General of Technical Development, from the various annual reports on the *Working and Administration of the Company Act*, 1956, Department of Company Affairs, Govt. of India; on gross block investment in the public sector undertakings from *Annual Report on the Working of Industrial and Commercial Undertakings of the Central Government*, Bureau of Public Enterprise, Govt. of India; on industrial licences issued to various states from *Uzair Committee Report on Industrial Licensing*, Planning Commission (1969) and *Assochem Parliamentary Digest* (No. 9, April 1984); data on tractors, oil engines and pumping sets used for irrigation were
obtained from the various rounds of quinquennial Livestock Census, Directorate of Economics and Statistics, Ministry of Agriculture; while fertilizer use data were obtained from the various issues of Fertilizer Statistics, Fertilizer Association of India. For working out values in per capita terms for non-census years, we used mid-year population projection obtained from the Report of the Expert Committee On Population Projection, Office of the Registrar General, Govt. of India (1978). All the data in monetary terms were suitably deflated by the relevant deflators, as given in Chandhok (1978), at 1970-71 prices. If data for the corresponding years were not available, we have taken the nearest possible information in terms of time to represent that year.