CHAPTER 5

DATA SOURCES AND MEASUREMENT OF VARIABLES

5.1 Introduction:

The importance of the data base in an empirical study like the present one needs hardly be emphasized. The longer the time series the greater the problem encountered in ensuring consistency of the data over time and consistency of the different series with each other. The main data source for the present study is the Annual Survey of Industries of the Central Statistical Organization. The data collected pertains to the organized or registered manufacturing sector or the factory sector. The data aggregated at the two digit level of classification is used in the study. The study covers eighteen industry groups at the two digit level and the period covered extends from 1960-61 to 1992-93. The nonparametric production frontier framework used for deriving capacity utilization measures requires data on output as well as four inputs - namely, capital labor, energy and intermediate materials.

In the section which follows i.e. Section 5.2, the nature of the data sources used in the study is reviewed. In Section 5.3 a description of the measurement of the variables used in the study is given while Section 5.4 contains some concluding remarks.
5.2 Sources of Data:

The Annual Survey of Industries (ASI) conducted every year since 1959 by the National Accounts Statistics Office and processed by the Central Statistical Organization (CSO) is the main source of data for this study. The ASI is an important data source for studies on the organized industrial sector in India as it provides a fairly consistent data set over a long period of time and at detailed levels of disaggregation. It provides detailed information on various characteristics like value added, value of output, employment, capital stock, wages, value of fuels consumed, value of intermediate inputs used etc.

The data from the ASI are available for the organized or registered manufacturing sector. The coverage of the ASI extends to the entire ‘Factory sector’ comprising of all industrial units called ‘factories’. The ASI data cover the factory sector with the exception of defence factories, oil storage depots and technical training institutes. The Factory sector includes both census and sample sectors. The census sector includes factories in which the manufacturing process is carried on with the aid of power and which employ on an average fifty or more persons, as well as those in which the manufacturing process is carried on without the aid of power and which employ 100 persons or more, and all electricity undertakings. The sample universe consists of the remaining factories registered under section 2(m)(i) and 2(m)(ii) of the Factories Act, 1948. These are units which employ between ten and ninety-nine workers without the aid of power.
The manufacturing sector is divided into 23 two-digit major groups which are further subdivided into three digit groups. In most cases information is provided at a still lower level of disaggregation.

The criterion for the classification of a factory in ASI is the value of its principal products. This often results in shifts of factories from one industrial class to another in subsequent surveys and thus affects the comparability of data over time.

Another problem with the ASI data was posed by the change in the industrial classification since 1973-74 when the National Industrial Classification replaced the ASI classification used since 1959. In addition, over time there has been a change in the reference period. The reference period for the ASI data from 1959 to 1965 is the calendar year for all industries except in cases of sugar, cotton ginning, cleaning and processing and electric light and power. From ASI - 1966 however the reference period is the accounting year of the factory ending on any day between 1st April and 31st March of that year.

Till 1971-72 the data were available separately for the census and sample sectors. Since the census and sample figures can be matched only at the three-digit level of disaggregation the data had to be first collected at the three digit level, and then aggregated upto the two digit level.

Despite certain weaknesses, the ASI data still remains the basic data source for detailed information on various economic characteristics.

In addition to the ASI data which is the principal data source, the current study also used data published in the Reserve Bank of India
Bulletin (RBI). The output series from ASI is at current prices. For the purpose of deflation the appropriate price indices for the different industries from the official series - Index Number of Wholesale Prices in India (base 1970-71) have been used. Again the value of fuels consumed and the value of intermediate materials used by the different industry groups as given in the ASI is in current prices. To obtain the values in real terms (with base 1970-71) the appropriate Wholesale Price Indices (WPI) have been used. The WPI of Fuels, Light and Lubricants is used for fuels consumed while for materials, an appropriate weighted price index of intermediate inputs is used. The detailed description of these indices used is given in Section 5.3.

Data on Unit Value Index of Materials (UVIM) have been obtained from various issues of the RBI Report on Currency and Finance. Three different series on UVIM with base years 1958, 1968-69 and 1978-79 are available. For the purpose of the study the three series were spliced to obtain a UVIM series for the entire period with a common base. This series was then converted to base 1970-71.

The price index of construction used in the study (for capital stock estimation) is the implicit deflator obtained as the ratio of Gross Domestic Capital Formation at current and constant (1970-71) prices from the National Account Statistics of the CSO.

### 5.3 Measurement of Variables:

The current study based on nonparametric linear programming approach to estimation of production frontier requires time series data on output and the four inputs namely, capital, labor, energy and
intermediate materials for the eighteen industry groups covered in the study. The section below discusses the measurement of the variables and the issues involved.

**Measurement of Output**: The data collected from the ASI on value of output is at current prices and therefore must be corrected for price changes. The appropriate Wholesale Price Indices with base 1970-71 from the RBI bulletin have been used for purposes of deflation to obtaining the value of output at constant prices. The detailed categories for which the wholesale price data are available do not match exactly with the two digit classification of the ASI. A close and detailed scrutiny of the available data was required before selecting the suitable price deflators.

Also while ASI data on the value of output produced is given at ex-factory price (which do not include excise duties and other distributive margins) it may not be entirely appropriate to use the WPI numbers (which excludes excise duties) to deflate it. The study has however used the WPI numbers for deflation purposes.

**Measurement of Capital Stock**: The problem of constructing a time series of capital in real terms has received a lot of attention; yet there is no consensus about a unique measure of real capital. Several theoretical and empirical problems are involved in measuring capital stock. Goldar (1986) provides a very useful review of both the conceptual problems and the shortcomings of the various existing estimates of capital stock for Indian manufacturing.

The first problem is to determine whether gross fixed capital formation or net fixed capital formation should be adopted as the measure of capital input. Ideally, for purposes of economic analysis it
is desirable to use the estimate of net capital stock provided a reasonable measure of true economic depreciation can be found out. But existing estimates of depreciation are either tax-based accounting concepts or based on certain rules of thumb.

The next problem relates to the measurement of a time series of real capital. Capital is usually measured by the "perpetual inventory method". In this method, the time series of the stock of capital is built up step by step from time series of (Rupee value of) investment and prices of capital goods.

For Indian industry most of the earlier studies used unsatisfactory measures of capital input. Banerji’s (1975) study while using an appropriate deflator for capital goods prices is based on an arbitrary assumption for obtaining base year capital stock. Hashim and Dadi’s (1973) study represents a significant improvement over earlier studies. In particular they have paid close attention towards obtaining the base year capital stock. The limitation however is in the capital goods deflator used - the use of price index of manufactured articles rather than a price index based on machinery and construction prices. Ahluwalia’s (1985) study also draws heavily on Hashim and Dadi’s estimates of gross fixed capital stock at replacement cost for 1960, but uses the wholesale price index for machinery and equipment for deflating gross investment series. By far the most detailed attention paid to the measurement of capital stock for Indian industry is the study by Goldar (1986). The study gives a detailed review of the various existing estimates for capital stock and the estimates provided by Goldar are a clear advancement over the earlier studies. Careful
attention is paid to obtaining the base year capital stock, obtaining an appropriate deflator and making allowances for discarding of assets.

For the purpose of the present study, the measure of gross fixed capital stock at constant prices is obtained as follows. The measure of capital stock includes land and excludes working capital. Working capital has been excluded in many earlier studies including Goldar's (1986). The estimates of capital stock are also gross of depreciation.

The perpetual inventory method is used to obtain the time series on capital stock. Let $K_0$ denote the base year capital stock, $I_t$ (1960) the gross investment (at base year prices) in fixed assets in year $t$, then fixed capital stock in year $T$ denoted by $K_T$ is given by:

$$K_T = K_0 + \sum_{t=0}^{T} I_t$$

The gross investment $I_t$ is given by:

$$I_t = \frac{[B_t - B_{t-1} + D_t]}{P_t}$$

where $B_t$ is the book value of fixed assets at the end of year $t$, $D_t$ is the amount of depreciation allowances made during year $t$ and $P_t$ is the capital goods price deflator.

The methodology in the present study for $I_t$ is thus similar to that used by Goldar (1986). The deflator used is also similar: a weighted average of price indices of construction and machinery is used, the weights being the relative magnitudes of these two categories of assets in the base year. For construction, the implicit price deflator computed as the ratio of the index of gross domestic capital formation at current and constant (1970-71) prices obtained
from the National Accounts Statistics (CSO) is used. For machinery and transport equipment, two price indices, one for machinery produced domestically and the other for imported machinery are combined. For domestic machinery, the official Wholesale Price Index Number of Machinery and Transport Equipment from the RBI is used. For imported machinery the unit value index of imports (UVIM) is used. For the period of the study, the unit value indices are available at three different base years, 1958, 1968-69 and 1978-79. From the available series, first a UVIM with common base is constructed for the entire time period and then converted to base 1970-71. To combine the two price indices, assumptions similar to those made by Goldar. Are made. Since the proportion of imported machinery in total investment in machinery has been declining, the two price indices (domestic and imported) have been combined in the ratio of 1:1 for the period 1960-65 and in the ratio 3:1 for the subsequent period.

To obtain the base year (1960) capital stock the study relies heavily on Hashim and Dadi’s study. The gross-net ratios given by Hashim and Dadi at the three digit level are used and assumptions similar to those made by them have been made for the purpose of the present study. For land, like Goldar the study assumes the gross-net ratio to be unity. From the ASI (1960) the book values of (1) land, (2) building and construction, (3) plant and machinery and (4) other assets, as existing at the end of 1960, are obtained. Multiplying these figures by their corresponding gross-net ratios, and then adding gross fixed assets at purchase prices for industries at the three digit level (e.g. S381) are obtained. From the Hashim-Dadi study, the ratio of gross
fixed assets at current (1960) prices and at purchase prices at the two digit industry level (e.g. Metal product $MP_{38}$) are obtained. This ratio is then multiplied by the figure of gross fixed assets at purchase prices obtained at the three-digit level (this three digit industry being a constituent part of the two digit industry whose ratio has been obtained). Thus, gross fixed assets in the three digit industry at the end of 1960 at current (1960) prices is obtained as ($K_{381}$):

$$K_{381} (1960) = S_{381} \times MP_{38}$$

This procedure thus enables the estimation of gross fixed assets at the end of 1960 at 1960 prices for all the eighteen two digit industry groups in the study. The base year capital stock is then converted to constant prices using the deflator constructed. To this base year stock (at constant 1970-71 prices), yearly deflated gross values of investment at constant (1970-71) prices were added to obtain the capital stock series at constant prices. In the present study no allowance has been made for discarding of assets.

Measurement of Labor Input: In the case of labor, the stock available to the industry is the number of persons employed by it during a year. The ASI publishes annual data on ‘workers’ as well as ‘employees’ and either of them can be used as a measure of labor input. ‘Total employees’ as a measure of labor input includes both ‘workers’ as well as ‘persons other than workers’. The latter category of employees consists of supervisors, technicians, managers, clerks and other similar types of employees. It has been argued that this category of employees is as important for getting the work done as the workers who operate the machines and therefore their services should be taken
into account in the measurement of labor input (see Sinha and Sawhney, 1970).

In the present study the ASI data on total employees is taken as a measure of labor input. Using total employees as a measure of labor input thus involves the assumption that ‘workers’ and ‘persons other than workers' are perfectly substitutable. This is clearly an assumption which is not realistic and is thus a limitation of this measure of stock of labor input.

**Measurement of Energy Input:** The ASI publishes annual data on total value of fuels consumed at current prices. Fuels consumed represents total value of all items of fuels, lubricants, electricity, water etc. consumed by the factory during the accounting year except those that directly enter into products as materials consumed. To obtain the value of fuels consumed at constant prices the price index of energy i.e. the official Wholesale Price Index Number of Fuels, Lights and Lubricants with base 1970-71 from the RBI Bulletin is used.

**Measurement of Materials Input:** Again the ASI is the source of published data on value of total materials consumed. These data are in current prices. Total materials consumed represents the total delivered value of all items of raw materials, chemicals, packing materials and stores which actually entered into the production process of the factory during the accounting year. To arrive at the value of intermediate goods consumed at constant prices the study uses a weighted constructed price index for materials which was derived for each industry separately. For each industry the ASI gives detailed quantity and expenditure data on a large number of items. The detailed data are available for five broad groups of intermediate inputs.
consumed: Basic materials, chemicals and auxiliary materials, packing materials, consumable stores and materials consumed for repair and maintenance. For the base year (1960) the shares of these five intermediate inputs consumed in total value of materials consumed was obtained. In most cases basic materials account for 80 to 90 per cent of the total expenditure on materials. Using these shares as weights a weighted price index of materials input was obtained. After going through the list of basic materials consumed by each industry the study has used appropriate indices from the official series on Index Number of Wholesale Prices. For Chemicals and Auxiliary materials the Wholesale Price Index Number of Chemicals and Chemical Products is used. For packing materials, a weighted average of Price Indices of Paper Products, Wood and Wood Products and Jute Hemp and Mesta Textiles is obtained. For consumable stores, again a weighted average of price indices of Wood and Wood Products, Paper and Paper products, Non-metallic Mineral Products, Basic Metals Alloys and Metal Products and Chemical and Chemical products is used. For materials consumed for repair and maintenance the price index of Machinery and Machine tools is used. For each industry, thus, a weighted price index of intermediate goods consumed is constructed. This index is then used to obtain the value of intermediate goods consumed at constant (1970-71) prices.

5.4 Conclusions:

Thus, the ASI is the primary source of data for the present study. Despite several weaknesses, it remains an important data source for
studies on the industrial sector. The present study has also paid particular attention to the measurement of the variables, especially to the measurement of capital stock and also to the construction of a deflator for intermediate materials input. Despite the attention paid to the measurement of the variables, it is important to keep in mind the limitations in the use of such a long time series for empirical purposes of obtaining trends in capacity utilization.