

Measuring Productivity at the Industry Level
THE INDIA KLEMS DATA BASE

DATA MANUAL 2021
(Version 8)

**Sadhan Kumar Chattopadhyay, Siddhartha Nath,
Sreerupa Sengupta and Shruti Joshi**

Reserve Bank of India

AUTHORS

Sadhan Kumar Chattopadhyay

Director
Reserve Bank of India
skchattopadhyay@rbi.org.in

Siddhartha Nath

Assistant Adviser
Reserve Bank of India
siddharthanath@rbi.org.in

Sreerupa Sengupta

Manager
Reserve Bank of India
ssengupta1@rbi.org.in

Shruti Joshi

Manager
Reserve Bank of India
sjoshi@rbi.org.in

The Data Manual has been prepared under the guidance of Professor Bishwanath Goldar, Professor Suresh Aggarwal, Dr. Abdul Erumban, and Shri. Pilu Chandra Das.

The Authors thank Shashank Bhujade for editorial support.

Contents

Chapter 1: Introduction	3
1.1 Background	3
1.2 Coverage: Industries and Variables.....	3
Chapter 2: Gross Value-Added Series at the Industry Level	8
2.1 Methodology	8
2.2 Implementation Procedure	8
Chapter 3: Gross Output Series at the Industry Level	11
3.1 Methodology	11
3.2 Implementation Procedure	12
Chapter 4: Labour Input Series at the Industry Level	15
Data Source, and Methodology	15
Chapter 5: Capital Input Series at the Industry Level	21
5.1 Methodology	21
5.2 Data and sources.....	22
5.3 Outstanding Issues.....	29
Chapter 6: Intermediate Input Series at the Industry Level	34
6.1 Methodology	34
6.2 Implementation Procedure	35
Chapter 7: Factor Income Share Series at the Industry Level	41
7.1 Methodology for Measuring Labour Income Share Series	41
7.2 Implementation Procedure	43
7.3. Outstanding Issues.....	48
Chapter 8: Growth Accounting Methodology	49
8.1 Methodology for Measuring Productivity Growth at the Industry level.....	49
8.2 Methodology for Aggregation across Industries	51

List of Appendixes

Appendix A: Definitions of Employment in NSSO employment & unemployment surveys .	17
Appendix B: Some Discussions on the Estimation of Capital Services.....	31

List of Tables

Table 1. 1: Industrial Classification for INDIA KLEMS Database 2021	5
Table 1. 2: Variables in our Multifactor Productivity Database for 27 Industries (Annual Time Series 1980-81 onwards).....	6
Table 2. 1: List of Manufacturing Industries for which Gross Value-Added data is obtained by adjusting data for NAS industries (From 1980 to 2011).....	9
Table 2. 2: List of Manufacturing Industries for which Gross Value-Added data is obtained by adjusting data for NAS Industries (2011 onwards).....	10
Table 3. 1: List of Manufacturing Industries for which Gross Output is directly available from NAS.....	12
Table 3. 2: List of Manufacturing Industries for which Gross Output is obtained by adjusting data for NAS Industries (2011 Onwards)	13
Table 5. 1: Capital asset types in National Accounts Statistics and corresponding assets in India KLEMS	24
Table 5. 2: Asset Types in ASI and India KLEMS.....	26
Table 5. 3: Depreciation Rate by Asset Type Used in the Computation of Capital Input.....	28
Table 7. 1: NAS Sectors and Corresponding Study Industries, for Computation of Labour Income Shares	42
Table 7. 2: Industries and Groups for which η (proportion of labour income out of mixed income) has been estimated	47

List of Box

Box 1: Natural gas price index.....	38
-------------------------------------	----

Chapter 1: Introduction

1.1 Background

This document describes the procedures, methodologies and approaches used in constructing the India KLEMS database version 2021. Till now, the work relating to compilation of India KLEMS data was being done by the India KLEMS team housed at the Centre for Development Economics, Delhi School of Economics. Now the work has been shifted to the KLEMS Division of Department of Economics and Policy Research of RBI since February 2021 and hence the KLEMS data have been compiled independently by the KLEMS Division under the supervision of India KLEMS team. This work is meant to support empirical research in the area of economic growth. In addition, the database is meant to support the conduct of policies aimed at supporting acceleration of productivity growth in the Indian economy, requiring comprehensive measurement tools to monitor and evaluate progress. Finally, the construction of the database would also support the systematic production of reliable statistics on growth and productivity using the methodologies of national accounts and input-output analysis.

Following the earlier trend, this version of database includes measures of economic growth, employment creation, capital formation and productivity at the industry level from 1980-81 onwards. The input measures will incorporate various categories of Capital (K), Labour (L), Energy (E), Materials (M) and Services (S) inputs.

The present document describes the India KLEMS database version 2021. In fact, the present version is an extended India KLEMS research project, “Disaggregate Industry Level Productivity Analysis for India- the KLEMS Approach” undertaken at the Centre for Development Economics, Delhi School of Economics.

The dataset includes measures of Gross Value Added (GVA), Gross Value of Output (GVO), Labour (L), Capital (K), Energy (E), Material (M), Services (S), Labour Quality (LQ), Labour Productivity (LP) and Total Factor Productivity (TFP) at the industry and economy level from 1980-81 onwards. The database covering the period 1980-81 to 2019-20 has been constructed on the basis of data compiled from NSO, NSSO, ASI, Input-Output tables (I-O tables) and processed according to appropriate procedures. These procedures were developed to ensure harmonization of the basic data, and to generate growth accounts in a consistent and uniform way. Harmonization of the basic data has focused on several areas such as industrial classification, and aggregation levels. The data base covers 27 industries comprising the entire Indian economy. The industries are shown in Table 1.1 below. The variables in the data set are given in Table 1.2.

1.2 Coverage: Industries and Variables

In this section we describe the coverage of the India KLEMS database in terms of industries and variables. In principle, the 39-year period from 1980-81 (1980) to 2019-20 (2019) is

covered. At a disaggregated level, database is created for 27 industries. The industrial classification is constructed by building concordance between NIC 2008, NIC 2004, NIC 1998, NIC 1987 and NIC 1970 to generate continuous time series from 1980 to 2019. This classification is very close to the International Standard Industrial Classification (ISIC) revision 3. The 27 industries are aggregated to form six broad sectors, namely:

- Agriculture, Hunting, Forestry and Fishing
- Mining and Quarrying
- Manufacturing
- Electricity, Gas and Water supply
- Construction
- Services

Table 1.1 provides a listing of the 27 industries, including the higher aggregates. For detailed classification and concordance of study industries with NICs, readers may refer to Version 7 of the manual published in 2021.

Table 1.2 provides an overview of all the series included in our database. Measures of Capital (K), Labour (L), Energy (E), Material (M) and Service (S) inputs as well as Gross Output (GO), have been constructed using National Accounts Statistics (NAS), Annual Survey of Industries (ASI), NSSO rounds and Input-Output (IO) Tables. In building annual time series on gross output, five inputs and factor income shares, various assumptions are made to fill up gaps in industry details and link series over time. As we know that NSSO rounds of unregistered manufacturing, Input Output Transaction Tables, and Employment and Unemployment Surveys by NSSO are available only for certain benchmark years. The use of information from these data sources necessitates interpolation and assumption of constant shares for building series of output and inputs. Further the construction of growth accounting series like total factor productivity, labour productivity, etc. are based on theoretical models of production and needs additional assumptions that are spelt out in subsequent chapters of the manual. Finally, the other Series like NDP at factor cost, compensation of employees, *etc.* are additional series which are used in generating the growth accounts and are informative by themselves.

Table 1. 1: Industrial Classification for INDIA KLEMS Database 2021

Sl. No.	Description of Industry
1	Agriculture, Hunting, Forestry and Fishing
2	Mining and Quarrying
3-15	MANUFACTURING SECTOR
3	Food Products, Beverages and Tobacco
4	Textiles, Textile Products, Leather and Footwear
5	Wood and Products of Wood
6	Pulp, Paper, Paper Products, Printing and Publishing
7	Coke, Refined Petroleum Products and Nuclear Fuel
8	Chemicals and Chemical Products
9	Rubber and Plastic Products
10	Other Non-Metallic Mineral Products
11	Basic Metals and Fabricated Metal Products
12	Machinery, nec
13	Electrical and Optical Equipment
14	Transport Equipment
15	Manufacturing, nec; recycling
16	Electricity, Gas and Water Supply
17	Construction
18-27	SERVICE SECTOR
18	Trade
19	Hotels and Restaurants
20	Transport and Storage
21	Post and Telecommunication
22	Financial Intermediation
23	Business Services
24	Public Administration and Defense; Compulsory Social Security
25	Education
26	Health and Social Work
27	Other Services

Source: Prepared by authors following EU KLEMS.

**Table 1. 2: Variables in our Multifactor Productivity Database for 27 Industries
(Annual Time Series 1980-81 onwards)**

Variable Description
GVA
Gross value added (GVA) at current prices
Gross value added (GVA) at constant prices
Annual growth rate in GVA (in per cent)
GVO
Gross value of output (GVO) at current prices
Gross value of output (GVO) at constant prices
Annual growth rate in GVO (in per cent)
Labour Input
Labour employment persons
Growth rate of labour employed (in per cent)
Labour quality index
Growth rate of labour quality index
Growth rate of labour input
Labour income share in GVA
Labour income share in GVO
Capital Input
Capital stock (K) at current price
Capital stock (K) at constant price
Capital Composition index
Growth rate of capital stock (in per cent)
Growth rate in capital services (in per cent)

Capital income share in GVA
Capital income share in GVO
Energy Input
Energy input series
Share of energy input in GVO
Material Input
Material input series
Share of material input in GVO
Service Input
Service input series
Share of service input in GVO
TFP (MFP)
Growth of total factor productivity (in per cent)

Chapter 2: Gross Value-Added Series at the Industry Level

This chapter describes the data sources and methodology used to construct the Gross Value Added (GVA) series at current and constant prices for 27 study industries for the period of 1980-81 (1980) to 2019-20 (2019).

2.1 Methodology

The National Accounts Statistics (NAS) brought out by the NSO (National Statistics Office, Government of India) is the basic source of data for the construction of series on gross value added for INDIA KLEMS-industries. Up to 2011-12, estimates of GVA at both current and constant (2011-12) prices for all industries are obtained from Back Series of National Accounts (Base 2011-12). For years after 2011-12, they are directly obtained from NAS 2022. However, NAS estimates of value added are not available for a few India KLEMS industry groups. Therefore, we had to split some of the aggregate industry groups from NAS. Maintaining consistency with NAS, the value-added data in India KLEMS is measured at basic price.

2.2 Implementation Procedure

The construction of Gross valued added series involves three steps.

Step 1: A concordance table between the classification used in the NAS and the 27 study industry classification used for this project has been prepared. Further, concordance between all the 27 sectors has been constructed with NIC- 1970, 1987, 1998, 2004 and 2008. Out of the 27 study industries, for 20 industries, GVA series both in current and constant prices is directly taken from NAS¹. The sectors for which data are provided in NAS are Agriculture, Forestry & logging, Fishing, Mining and Quarrying, Manufacturing, Electricity, gas and water supply, Construction, Trade, Hotels & Restaurants, Railways, Transport by other means, Storage, Communication, Banking & insurance, Real estate, Ownership of Dwelling & Business Services, Public Administration & Defense and Other Services.

Step 2: For manufacturing industries where direct estimates of GVA were not available from NAS, estimates have been made using additional information from ASI and NSSO unorganized manufacturing data. In these cases, GVA data is constructed by splitting the NAS data using ASI or NSSO distributions. ASI data (annual) has been used for registered and corporate manufacturing whereas interpolated ratios from NSSO 40th (1984-85), 45th (1989-90), 51st (1994-95), 56th (2000-01) 62nd (2005-06), 67th round (2010-11) and 73rd round (2015-16) rounds have been used for

¹ From both aggregated 9 NAS sectors Gross Domestic Product by economic activity statement along with the disaggregated statements of these 9 NAS sectors.

Unregistered and household manufacturing segments. For unregistered manufacturing for years 2015-16 to 2019-20 the ratio obtained from 73rd round of unregistered manufacturing has been used for the splitting. Table 2.1 and 2.2 showcases the methodology used to split GVA of certain NAS sectors to match concordance with our classification.

Table 2. 1: List of Manufacturing Industries for which Gross Value-Added data is obtained by adjusting data for NAS Industries (from 1980 to 2011)

Industry No.	Industry description	NAS industry description	Methodology
5	Wood and products of Wood (20)	Wood and Wood Products, Furniture, Fixtures <i>etc.</i> (20+361)	From 1980 to 2011 Gross Value Added of NAS sector (20+361) is split using ASI and NSSO distributions.
7	Coke, Refined Petroleum and Nuclear Fuel (23)	Rubber, Petroleum Products (23+25)	Used ASI and NSSO proportions to split GVA of NAS sector (23+25) into separate 23 and 25. Since 2011-12 NAS provides separate series for 23 and 25.
9	Rubber and Plastic Products (25)	Rubber, Petroleum Products (23+25)	Same as above
11	Basic Metals and Fabricated Metal Products (27+28)	Basic Metals (271+272+2731+2732)	Use ASI and NSSO proportions to split 28, 29 and 30.
		Metal Products and Machinery (28+29+30)	Add fraction of 28 to Basic metals (27) obtained from NAS
12	Machinery, nec. (29)	Metal Products and Machinery (28+29+30)	Use ASI data to split GVA of NAS sector Metal products and machinery(28+29+30) to separate 29 from 28+29+30
13	Electrical and Optical Equipment (30 to 33)	Electrical Machinery (31+32)	Split GVA of NAS sector Metal Products and machinery (28+29+30) using ASI proportions.
			Add fraction of 30 to Electrical machinery (31+32) obtained from NAS
15	Manufacturing nec; Recycling (36+37)	Other Manufacturing (33 +369)	Split GVA of NAS sector Other manufacturing (33+369) using ASI proportions.
		Recycling (371+372)	Add fraction of 36 and 361 to Recycling (371+372)

Note: The figures in parentheses are two-digit NIC 2004 codes

For details of the method refer KLEMS Data Manual Sept 2021 release

Source: National Account Statistics (NAS), Annual Survey of Industries (ASI) and National Sample Survey Office

Table 2. 2: List of Manufacturing Industries for which Gross Value-Added data is obtained by adjusting data for NAS Industries (2011 Onwards)

Industry No.	Industry description	NAS industry description	Methodology
11	Basic Metals and Fabricated Metal Products (24+251+259)	Manufacture of basic iron and steel (241+2431) + Manufacture of basic precious and non-ferrous metals (242+2432) + Manufacture of fabricated metal products, except machinery and equipment (25)	Used ASI and NSSO proportions to split GVA of NAS sector (25) into separate 251, 252 and 259
12	Machinery, nec. (28+252+275+304+3311+3312)	Manufacture of fabricated metal products, except machinery and equipment (25) + Manufacture of electrical equipment (27) + Manufacture of machinery and equipment n.e.c (28) + Manufacture of transport equipment (29+30) + Repair and installation of machinery and equipment (33)	Used ASI and NSSO proportions to split GVA of NAS sector (25, 27, 29+30, 33)
13	Electrical and Optical Equipment (26+27-275+325+3313+3314+3319+3320)	Manufacture of computer, electronic & optical products (26) + Manufacture of electrical equipment (27) + Other manufacturing (32) + Repair and installation of machinery and equipment (33)	Used ASI and NSSO proportions to split GVA of NAS sector (27, 32 & 33)
14	Transport Equipment (29+30-304+3315)	Manufacture of transport equipment (29+30) + Repair and installation of machinery and equipment (33)	Used ASI and NSSO proportions to split GVA of NAS sector (29+30) into separate 29+30-304 & 304.
15	Manufacturing, nec; recycling (31+32-325)	Manufacture of furniture (31) + Other manufacturing (32)	Used ASI and NSSO proportions to split GVA of NAS sector (32) into separate 32-325 & 325.

Note: The figures in parentheses are two-digit NIC 2008 codes

For details of the method refer KLEMS Data Manual Sept 2021 release

Source: National Account Statistics (NAS), Annual Survey of Industries (ASI) and National Sample Survey Office

It is important to note that the industry level value added volume indices are based on NAS. NSO provides single deflated value-added estimates for all sectors except Agriculture.

Step 3: According to India KLEMS, GVA is adjusted for Financial Intermediation Services Indirectly Measured (FISIM). The details method of FISIM adjustment is provided in earlier versions of KLEMS manual.

Chapter 3: Gross Output Series at the Industry Level

This chapter describes the procedures and methodologies used in constructing the database for gross output series at the industry level over the period 1980-81(1980) to 2019-20(2019).

3.1 Methodology

To construct the gross output series at industry level, we use multiple data sources namely National Accounts Statistics, Annual Survey of Industries, NSSO rounds for unorganized manufacturing and Input Output Transaction tables. The data sources and methodology used are documented below:

National Accounts Statistics:

The NAS is the basic source of data for the construction of time series on the gross output. The NAS back series 2011 with base 2004-05 and NAS 2014 provides estimates of gross output for six disaggregate industries at current and constant prices since 1950-51 till 2011-12. These sectors are Agriculture, Mining and Quarrying, Construction and Manufacturing sectors (Registered and Unregistered Manufacturing). However, the Back Series with base 2011-12, which is the source of GVA in India KLEMS database 2021, does not provide estimates of GVO for most of the industries except for Agriculture, Mining and Quarrying and Construction. For these three industries GVO data at current and constant prices directly obtained from Back Series with base 2011-12. Therefore, for 1980-81 to 2011-12 period, we estimate the GVO series for remaining 24 industries at current and constant prices by applying the respective GVO/GVA ratio for current and constant prices obtained from Back Series with base 2004-05 and NAS 2014 to GVA with 2011-12 base. For years since 2011, we take the estimates of GVO both at current and constant prices for all industries directly from NAS 2022.

(a) Filling procedures of National Accounts series: It is to be noted that the NAS estimates of gross output for a few industry groups are at a more aggregate level, requiring splitting of the aggregates. In such cases, NAS estimates of output have been split using additional information from Annual Survey of Industries and NSSO rounds of Unregistered Manufacturing to obtain estimates at higher level of disaggregation. Secondly, for Unregistered manufacturing gross output data is available in NAS from 2004-05 onwards. In this case, information from NSSO survey rounds has been used for missing years to derive output estimates of unregistered manufacturing industries at current and constant prices.

As mentioned earlier, for gross value-added series of service sectors we obtain our estimates from NAS. However, prior to 2011-12 National Accounts do not provide any estimates of gross output

of service sectors and hence we rely on Input output transaction tables (from which the ratio of gross to value added is computed which is then applied to GVA reported in NAS) which are available at an interval of 5 years or so. This necessitates interpolation and assumption of constant shares for measuring output of services sectors. The Input Output Transaction Tables for Benchmark years of 1978-79, 1983-84, 1989-90, 1993-94, 1998-99, 2003-04 and 2007-08 are used to derive gross output series for service sectors.

3.2 Implementation Procedure

The construction of the gross output series at current and constant prices involves the following steps:

Step 1: Measuring Gross Output of Agricultural Sector, Mining and Quarrying, and Construction

NAS provides nominal and real GVO series for a) Crops and Plantation, b) Animal Husbandry c) Forestry and Logging d) Fishing. By aggregating the GVO of these four subsectors we derive the GVO of Agricultural sector. The Gross output estimates of Mining and Quarrying and Construction at current and constant prices from 1980-2019 is also directly taken from NAS.

Step 2: Measuring Gross Output of Manufacturing Industries

Since 2011-12, gross output data for 7 out of 13 manufacturing industries listed in table 3.1 are directly picked up from NAS.

Table 3. 1: List of Manufacturing Industries for which Gross Output is directly available from NAS

Industry No.	Industry Description
3	Food Products, Beverages and Tobacco
4	Textiles, Textile Products, Leather and Footwear
5	Wood and Products of wood
6	Pulp, Paper, Paper products, Printing and Publishing
7	Coke, Refined Petroleum Products and Nuclear fuel
8	Chemicals and Chemical Products
9	Rubber and Plastic Products

Source: National Account Statistics (NAS) reports.

For the remaining 6 sectors output is constructed by splitting the NAS output data using ASI or NSSO distributions. ASI data (annual) has been used for registered manufacturing whereas

interpolated ratios from 67th (2010-11) and 73rd (2015-16) rounds have been used for Unregistered Manufacturing segments. A list of study industries is presented in Table 3.2 showcasing the methodology used to split GVO of certain NAS sectors to match concordance with our classification for the year 2011 to 2019.

Table 3. 2: List of Manufacturing Industries for which Gross Output is obtained by adjusting data for NAS Industries (2011 onwards)

Industry No.	Industry description	NAS industry description	Methodology
11	Basic Metals and Fabricated Metal Products (24+251+259)	Manufacture of basic iron and steel (241+2431) + Manufacture of basic precious and non-ferrous metals (242+2432) + Manufacture of fabricated metal products, except machinery and equipment (25)	Used ASI and NSSO proportions to split GVO of NAS sector (25) into separate 251, 252 and 259
12	Machinery, nec. (28+252+275+304+3311+3312)	Manufacture of fabricated metal products, except machinery and equipment (25) + Manufacture of electrical equipment (27) + Manufacture of machinery and equipment n.e.c (28) + Manufacture of transport equipment (29+30) + Repair and installation of machinery and equipment (33)	Used ASI and NSSO proportions to split GVO of NAS sector (25, 27, 29+30, 33)
13	Electrical and Optical Equipment (26+27-275+325+3313+3314+3319+3320)	Manufacture of computer, electronic & optical products (26) + Manufacture of electrical equipment (27) + Other manufacturing (32) + Repair and installation of machinery and equipment (33)	Used ASI and NSSO proportions to split GVO of NAS sector (27, 32 & 33)
14	Transport Equipment (29+30-304+3315)	Manufacture of transport equipment (29+30) + Repair and installation of machinery and equipment (33)	Used ASI and NSSO proportions to split GVO of NAS sector (29+30) into separate 29+30-304 & 304.
15	Manufacturing, nec; recycling (31+32-325)	Manufacture of furniture (31) + Other manufacturing (32)	Used ASI and NSSO proportions to split GVO of NAS sector (32) into separate 32-325 & 325.

Note: The figures in parentheses are two-digit NIC 2008 codes

For details of the method refer to KLEMS Data Manual, Version 7 released in Sept 2021

Source: National Account Statistics (NAS), Annual Survey of Industries (ASI) and National Sample Survey Office

Step 3: Measuring Gross Output for Services Sectors and Electricity, Gas and water supply

Since 2011 -12 NAS provided estimates of GVO at current and constant prices. Prior to 2011-12 Gross Output series for Services sectors and sector Electricity, Gas and Water supply have been constructed using information from Input – Output Transaction Tables of the Indian economy published by CSO. The details method of construction of GVO back series for services is provided in earlier versions of KLEMS Data manual.

Chapter 4: Labour Input Series at the Industry Level

This chapter provides information on the sources of data and method of measuring labour services. The aim is to estimate labour input so that it reflects the actual changes in the quantity (number of persons) and quality of labour input over time.

Data Source, and Methodology

The section discusses the construction of labour input for 27 industries from 1980 to 2019-20. Labour input is measured by combining data on labour persons and data on labour quality which is measured in terms of human capital embodied in the labour. In our study, we measure this on the basis of the educational qualification of the worker. It is important to segregate labor employed from labor quality because the contribution to output by each person also comes from this embodied capital. Moreover, the reward (wages and earnings) to each person also includes the reward for investment in human capital. Therefore, it is essential to separate out these differences in labour to clearly understand the underlying differences in labour characteristics. It is in this context that an endeavor has been made to estimate labour quality index. However, given the limitations of India's employment statistics (Sivasubramonian; 2004, & Himanshu; 2011, Ghose; 2016), especially the availability of information on wages/ earnings of different category of workers which could be used as an indication of their differences in ability makes it difficult to quantify these changes in the labour force in a pertinent way. Therefore, the study has computed the labour quality index based on five different education categories only.

This study aims to build a time series of employment and labor quality series for 27 industrial sectors. In this section we outline the sources of data and the methodology used. The section develops and implements the methodology of estimating persons employed (employment) and labour quality and combining the two to obtain the indices of labour input.

Sources of data

The Employment and Unemployment Surveys (EUS) of major rounds from 38th round (1983) to 68th round (2011-12) by National Sample Survey Office (NSSO) and the Periodic Labour Force Surveys (PLFS) of 2017-18, 2018-19 and 2019-20 by National Statistical Organization (NSO) are the main sources for estimating the total workforce in the country by industry groups, as per the National Industrial Classification (NIC). This year the study has adjusted the employment estimates obtained from Economic-Survey 2021-22 (pp.371).

Methodology

The employment in India has been computed in a step-wise manner as follows:

As a first step concordance between the KLEMS 27 sectors, and NIC-1970, 1987, 1998 and 2008 was done. Thereafter, the following steps were done:

1. KLEMS Industry wise distribution of workers by UPSS² across the four components - Rural Male, Rural Female, Urban Male and Urban Female were obtained from PLFS 2019-20. These numbers were then adjusted to the employment numbers given in the Economic Survey 2021-22 to find out the number of persons employed in the four segments and the total employment.
2. The 27-industry distribution of Employment from EUS is applied to the number of persons employed in step I to obtain L_{ij} for each industry where $i=1$ for rural and 2 for urban sectors, and $j=1$ for male and 2 for female for each round
3. The estimates of employment for the intervening years between the major EUS rounds and PLFS 2019-20 have been obtained by simple interpolation across the four segments- rural male, rural females, urban male and the urban females. After ignoring the outliers, the two closest rounds have then been used for interpolation. Further, the mid-point of the survey is around the mid-year of the survey rounds, i.e July for 38th round and 1st Jan for the other rounds; whereas the interpolation has been centered around 1st October which is the mid-year of the national income estimates³. Hence, the base is shifted to October of each year while interpolating⁴ the values. This is done to coincide the survey year mid-point to the mid-point of financial year.
4. Total persons in a year were then obtained for each industry as the sum of the L_{ij} over gender and sectors, i.e., $\sum_i \sum_j L_{ij}$.

For extrapolation backward to 1980-81 to 1982-83, the interpolation of the broad industrial classification of 32nd round (1977-78) and 38th round (1983-84) is used. Thus, the estimates from 32nd round are mainly used as control numbers. Between 2005 and 2011, we observed very high growth rate in number of employed persons series for the industry *Electrical and Optical Equipment* as compared to other manufacturing industries and the overall growth in employment in this industry during 2005-2011 was found to be higher than that for *Construction* (which seems somewhat unrealistic). It was also observed that there was a negative growth in employment in *Textiles, Textile Products, Leather and Footwear* although real GVA of this industry more than doubled between 2005 and 2011. To address these issues, some adjustments to the initial employment estimates for *Electrical & Optical Equipment* and *Textiles, Textile Products, Leather and Footwear* have been done. Onwards

² In the EUS, the persons employed are classified on the basis of their activity status into usual principal status (UPS), usual principal and subsidiary status (UPSS), current weekly status (CWS) and current daily status (CDS) for quinquennial rounds. refer to Appendix A for detailed definition of UPSS, CWS and CDS.

³ This is also the mid-point of financial year

⁴ For more detailed discussion on extrapolation and interpolation of the series, readers may refer to KLMES data manual version 2020 released by RBI on its website

2005, employment series for *Electrical & Optical Equipment* has been estimated applying annual growth rates obtained from ASI and NSSO rounds to the number of employed persons for the year 2005-06. Then, for the years 2005-06 to 2011-12, we compute the difference between estimated employment series from EUS rounds and that based on ASI & NSSO rounds for the *Electrical and Optical Equipment* industry and add it to *Textiles, Textile Products, Leather and Footwear industry*. This ensures that for the manufacturing as a whole our estimates remain the same as obtained from EUS rounds.

(b) Measuring quality index

The quality of labour force is of considerable importance in the context of productivity measurement, and one of the widely used methodologies to capture changes in labour quality is given by Jorgenson, Gollop and Fraumeni (JGF) (1987). In growth accounting methodology of measurement of total factor productivity (TFP) when output growth is decomposed into growth of inputs and the residual TFP, then labour input is measured as an index of labour service flows. It accounts for changes in labour quality in terms of labour characteristics such as educational attainment, age (experience), gender, employment status, etc. and thus accounts for heterogeneity of the labour force. In our study, the index of aggregate labour quality measures the changes in the composition of labour in the economy.

The present study has computed labour quality index by using only education characteristic in the JGF methodology. Thus, the data required for the labour quality index in the present case, is employment and earnings by education and by industry. The labour quality index has been computed using five education categories namely- up to primary, primary, middle, secondary & higher secondary, and above higher secondary. There are thus five types of persons employed for each of the 27 study industries. The quality growth rates are estimated for total persons employed in these industries in India for the 38th, 50th, 55th, 61st, 68th and PLFS rounds of NSO. They are then indexed to 1983 (38th round) as the base, so as to assess the temporal changes in labour's skill and interpolation has been done between the major rounds for values of the intermediate years. Since the series is required from 1980-81, we have extrapolated it backwards from 1983 and the index is recomputed with base 1980-81 equal to 100.

Therefore, the following steps have been performed:

1. Compute the distribution of persons employed by the five educational groups for all the 27 industries for the selected major rounds of EUS and PLFS.
2. Apply these proportions to the number of employed persons in different industries in the major rounds and PLFS to obtain the distribution of persons by education groups.
3. Estimate the earnings data from NSO which relates to mainly the regular and casual persons employed. It may however be mentioned that even for these two groups, for a large

number of persons employed, the wages are either missing or given as zero. For these employed persons whose wage information is missing, a Mincer equation is used to estimate the earnings for self-employed persons, using⁵ a Mincer wage equation, which is corrected for sample selection bias by using the Heckman's⁶ two step procedure. The Mincer function has been applied to the earnings of casual and regular employees where the earnings have been regressed on the dummies of age, gender, education, location (rural or urban), marital status, social exclusion and industry. The corresponding earnings of the self-employed are obtained as the predicted value with similar traits. The average wages per day are then computed for persons employed of different type of employment, i.e., self-employed, regular and casual combined together; whose wages are available.

4. Based on steps 1 to 3, labour quality index is estimated as equation 4.1

$$\begin{aligned} & \Delta L_{j,t} - \Delta \ln H_{j,t} & (4.1) \\ & = \sum \bar{v}_{l,j,t} \Delta \ln H_{l,j,t} - \Delta \ln \sum H_{l,j,t} \end{aligned}$$

Where; $\Delta \ln L_{j,t} = \sum \bar{v}_{l,j,t} \Delta \ln H_{l,j,t}$ with weights given by $v_{l,j}^L = \frac{P_{l,j}^L \cdot H_{l,j}}{\sum P_{l,j}^L \cdot H_{l,j}}$

$$\bar{v}_{l,j}^L = 0.5 \left[v_{l,j}^L t + v_{l,j}^L t - 1 \right]; \text{ and } \Delta \ln H_{j,t} = \Delta \ln \sum H_{l,j,t}$$

$\Delta \ln L_{l,j}$ indicates the growth rate (measured as log changes) of persons worked by labour type 'l' for industry 'j', $v_{l,j}^L$ is the share of worker type l in total labor compensation, measured as the product of wage rate (P) and employment (H). These weights are used as an average of period t and t-1.

$\Delta \ln \sum H_{l,j,t}$ is defined as the growth in employment where $\sum_l H_{l,j,t}$, indicate total persons worked in industry 'j' by all types 'l' in period 't'.

The first term of equation 4.1 indicates the changes in labour quality index, which is due to the changes in the composition of workers and the second term indicates the change in total persons employed in sector 'j'. The index of aggregate labour quality thus measures the changes in the composition of labour in the economy

⁵ In PLFS, the earnings by self-employed are also provided, but for consistency with other rounds we have used the same procedure of Mincer equation even for this round. In EU KLEMS (Timmer, et.al., op cit; p 67) it is assumed that the earnings of the self-employed is equal to the earnings of 'regular' employees.

⁶ For detailed explanation of Heckman two-step process one may refer to the previous year manual

The labour input (adjusted labour persons) then can be obtained by multiplying the number of persons employed by the corresponding labour quality index and the labour input growth is finally obtained by combining the growth of persons employed and the growth in the index of labour quality.

Appendix A: Definitions of Employment in NSSO employment & unemployment surveys

The surveys of NSSO on employment and unemployment (EUS) aim to measure the extent of ‘employment’ and ‘unemployment’ in quantitative terms disaggregated by various household and population characteristics following the three reference periods of (i) one year, (ii) one week, and (iii) each day of the week. Based on these three reference periods three different measures, termed as *usual status*, *current weekly status*, and the *current daily status*, are arrived at. While all these three approaches are used for collection of data on employment and unemployment in the quinquennial surveys, the first two approaches only are used for the purpose in the annual surveys.

Usual principal status: In NSS 27th round, the usual principal activity category of the persons was determined by considering the normal working pattern, *i.e.*, the activity pursued by them over a long period in the past and which was likely to continue in the future. For the identification of the usual principal status of an individual based on the major time criterion, in NSS 27th, 32nd, 38th, 43rd rounds, a trichotomous classification of the population was followed, that is, a person was classified into one of the three broad groups ‘employed’, ‘unemployed’ and ‘out of labour force’ based on the major time criterion. From NSS 50th round onwards, the procedure was changed and the prescribed procedure was a two-stage dichotomous one which involved a classification into ‘labour force’ and ‘out labour force’ in the first stage, and thereafter, the labour force into ‘employed’ and ‘unemployed’ in the second stage.

Usual subsidiary status: In the usual status approach, besides principal status, information in respect of subsidiary economic status of an individual was collected in all employment and unemployment surveys. For deciding the subsidiary economic status of an individual, no minimum number of days of work during the last 365 days was mentioned prior to NSS 61st round. In NSS 61st round, a minimum of 30 days of work, among other things, during the last 365 days, was considered necessary for classification as usual subsidiary economic activity of an individual.

Current weekly status: It is important to note at the beginning that in the EUS of NSSO, a person is considered as worker if he/she has performed any economic activity at least for one hour on any day of the reference week and uses the priority criteria in assigning work activity status. This definition is consistent with the ILO convention and used by most of the countries in the world for their labour force surveys. In NSSO, prior to NSS 50th round and in all the annual surveys till NSS 59th round, data on employment and unemployment in the CWS approach was collected by putting

a single-shot question ‘whether worked for at least one hour on any day during the last 7 days preceding the date of survey’. The information so collected was used to determine the CWS of the individuals. This procedure was criticized for being not able to identify the entire workforce, particularly among the women. It was then decided to derive the CWS of a person from the time disposition of the household members for the 7 days preceding the date of survey. The procedure was used for the first time in NSS 50th round. It is seen that the change in the method of determining the current weekly activity had resulted in increasing the WPR in current weekly status approach - more so for the females in both rural and urban areas than for males. The trend observed in NSS 50th round in respect of the WPR according to CWS suggested continuing with the procedure for data collection in CWS in NSS 55th and NSS 61st rounds.

Current Daily Status

Current Daily Status (CDS) rates are used for studying intensity of work. These are computed on the basis of the information on employment and unemployment recorded for the 14 half days of the reference week. The employment statuses during the seven days are recorded in terms of half or full intensities. An hour or more but less than four hours is taken as half intensity and four hours or more is taken as full intensity. An advantage of this approach was that it was based on more complete information; it embodied the time utilisation, and did not accord priority to labour force over outside the labour force or work over unemployment, except in marginal cases. A disadvantage was that it related to person-days, not persons. Hence it had to be used with some caution.

Chapter 5: Capital Input Series at the Industry Level

This chapter outlines the methodology employed to estimate capital input for the 27 industries in the India KLEMS database version 2022. All the productivity calculations in India KLEMS use capital services as an input to production, which is estimated using the approach developed by Jorgenson and Griliches (1967), and outlined in Jorgenson, Gollop, and Fraumeni (1987). The estimation of capital service growth rates is accomplished by first estimating capital stock for different capital assets, and then aggregating across assets after correcting for the differences in their marginal productivities. Capital services account for the differences in marginal productivity between different asset types while capital stock measures only the vintages of capital, they do not account for asset heterogeneity. For instance, aggregate capital stock measures would assume that a computer's productivity is the same as that of a car. However, proper measures of capital stock would account for the decline in the efficiency of both computers and cars over their lifetime. Following an overview of the theoretical method to measure capital services (Jorgenson and Griliches, 1967; JGF, 1987), the chapter discusses the specific empirical approaches we follow to implement these methods within the constraints of data availability for Indian industries. In India KLEMS, we have reported capital stock, capital services and the capital “capital composition effect”. Capital stock is reported in constant price volumes, and growth rates, while the latter two are reported in growth rates and indices.

5.1 Methodology

To measure capital services, using the Jorgenson and Griliches (1967) approach, we need estimates of capital stock for detailed asset types and the shares of each of these assets in total capital remuneration. The aggregate capital services growth rate is derived as a weighted growth rate of individual capital assets, the weights being the shares of each asset in the total compensation made to capital, *i.e.*,

$$\Delta \ln K_t = \sum_k \bar{v}_{k,t} \Delta \ln S_{k,t} \quad (5.1)$$

Where $\Delta \ln K_t$ is the aggregate capital service growth rate, $\Delta \ln S_{k,t}$ is the volume growth of capital asset k – or the growth rate of capital stock of asset k , and the weights $\bar{v}_{k,t}$ are the average shares of each asset in the value of total capital compensation in two periods such that the sum of shares over all capital types add to unity, *i.e.*,

$$\bar{v}_{k,t} = 0.5 \times (v_{k,t} + v_{k,t-1}), \text{ and } v_{k,t} = P_{k,t}^K S_{k,t} / \sum_k P_{k,t}^K S_{k,t} \quad (5.2)$$

where $P_{k,t}^K$ is the rental or user cost of capital asset type k in year t . Therefore, the numerator, which is the product of user cost of asset k and the capital stock of asset k , is the capital compensation received by asset k . The numerator is the sum of compensation received by assets, or the total capital compensation, and therefore $v_{k,t}$ is the share of asset k in total capital compensation. Under the neoclassical assumption of price-marginal product equality, it effectively incorporates the productivity or qualitative differences in the contribution of various asset types to aggregate capital services, as the composition of aggregate capital stock in any industry changes (see Jorgenson, 2001). For instance, as the marginal productivity of ICT capital is higher than that of other assets, a change in the composition of capital towards ICT capital will result in higher capital services, which will be captured by a larger value of the v for ICT assets.

It is evident from (5.2) that two important components of capital service measure are the asset wise capital stock, $S_{k,t}$ and the service price (rental price) of capital assets, $P_{k,t}^K$. Assuming a geometric depreciation rate, δ_k which is constant over time, but different for each asset type, capital stock in asset k in year t can be constructed using Perpetual Inventory Method (PIM) as:⁷

$$S_{k,t} = S_{k,t-1}(1 - \delta_k) + I_{k,t} \quad (5.3)$$

where, $I_{k,t}$ is the real investment in asset type k in period t . For a detail discussion on $P_{k,t}^K$, please see Box 5.1A in the Appendix B.

Once we have the estimates of capital stock for each asset (equation 5.3) and the growth rates of aggregate capital services using equation (5.1), we can calculate the capital composition effect ($\Delta \ln KQ_t$) - or the effect of changing asset composition of capital stock. This can be obtained as $\Delta \ln KQ_t = \Delta \ln K_t - \Delta \ln S_t$, where $S_t = \sum_k S_{k,t}$. These growth rates can then be indexed to the base year to obtain a capital composition index (KQ_t), i.e. $KQ_t = KQ_{t-1} \times e^{(\Delta \ln K_t - \Delta \ln S_t)}$ for $t >$ base year, and $KQ_t = 100$ for $t =$ base year. In the India KLEMS database, the capital composition growth rates are indexed to the first year, 1980.

5.2 Data and sources

Since our measure of capital input takes account of asset heterogeneity, it was essential to obtain investment data by asset type. We distinguish between 3 different asset types – construction, transport equipment, and machinery (includes ICT and non-ICT machinery).⁸ We exploit multiple

⁷ The estimates of capital stock using PIM are obtained at constant prices and are hence volume measures. Each asset's current price capital stock may be obtained by multiplying real capital stock with investment price deflators for the respective asset type. Then the aggregate capital stock in current prices ($P^I S_t$) can be obtained as:

$P^I S_t = \sum_k P_{k,t}^I S_{k,t}$, where $P_{k,t}^I$ is the investment deflator for asset k , $S_{k,t}$ is the capital stock in asset k .

⁸ This version of the database does not make a distinction between ICT and non-ICT assets, as the industry level data on ICT assets are weak. An attempt to estimate aggregate economy level ICT capital can be found in

sources of information for the construction of our database on capital services. This includes the National Accounts Statistics (NAS) that provide information on broad sectors of the economy, the Annual Survey of Industries (ASI) covering the organized manufacturing sector, the National Sample Survey Office (NSSO) rounds for unorganized manufacturing and various Input-Output tables. Even though we use multiple sources of data, our final estimates are fully consistent with the aggregate data obtained from the NAS. In addition, our approach to capital measurement is consistent with international practices such as the EU KLEMS⁹, which ensures the possibility of international comparisons. In what follows, we discuss the various sources of data for asset wise investment and the construction of the relevant variables, in detail.

(a) *Asset-wise investment for broad sectors of the economy*

Industry-level estimates of capital input require detailed asset-by-industry investment matrices. NAS provides information on aggregate capital formation by industry of use for nine broad sectors, which, nevertheless, was not sufficient for our purpose. Therefore, we have collected more detailed data on assets and industries from the CSO.¹⁰ This is the data underlying the published aggregate gross fixed capital formation by the broad industry groups, separately for public and private sectors. For those sectors for which the investment matrices were not available from CSO, we gather information from other sources (*e.g.* ASI for organized manufacturing and NSSO surveys for unorganized manufacturing) and benchmark it to the aggregate investment series from the National Accounts. The data used in the current version of the India KLEMS is based on the revised NAS with 2011-2012 base and is available only since 2012. Therefore, for earlier years, we extrapolate the series using growth rates from previous version of the data. Table 5.1 provides an overview of asset types available in NAS and their corresponding asset types used in our study. Investment in education and health are obtained directly from national accounts for the period after 2012, for each asset. For years before 2012, we assume the trend in the distribution of output, in order to split the total investment in the aggregates of these sectors into sub-sectors.

Erumban and Das (2015).

⁹ See O' Mahony and Timmer (2009) for a description of EU KLEMS database

¹⁰ This data is not publicly available. However, CSO has been kind to compile this data for the India-KLEMS project.

Table 5.1: Capital asset types in National Accounts Statistics and corresponding assets in India KLEMS

NAS Asset Types	India KLEMS Asset Types
Public Sector	
Buildings*	Construction
Other construction*	Construction
Transport Equipment ¹¹	Transport Equipment
Machinery & Equipment	Machinery & Equipment
Software (1999-00 onwards)	Machinery & Equipment**
ICT equipment (2012 onwards)	Machinery & Equipment**
Private Sector	
Residential buildings	Construction
Non-residential building	Construction
Other construction	Construction
Machinery & Equipment (incl. transport equipment)	Machinery & Equipment (transport equipment is excluded later*)
Software (1999-00 onwards)	Machinery & Equipment**
ICT equipment (2012 onwards)	Machinery & Equipment**

Source: CSO, NAS Different Issues

Note: The asset categories listed in this table are available for the period 1950-2016, for which we obtained detailed asset-industry data from the CSO. For years after 2016, when we relied on the published version of national accounts, the asset groupings are limited to dwellings, other buildings & structures, machinery & equipment (including transport equipment), intellectual property, and cultivated biological resources. Detailed data on intangible investment, such as intellectual property rights, R&D, and mineral exploration and evaluation, are also available in the 2012 base NAS data for the period 2011-2016. These are included in machinery assets in the KLEMS asset group, as we are unable to separate them from earlier versions of the data. * Buildings and other construction include dwellings, non-residential buildings, roads & bridges, and other construction. ** Software and ICT are included in the machinery and equipment, as we are not distinguishing between ICT and non-ICT assets in this version. *** Transport equipment was not separately available for the private sector, and therefore, it was imputed and subtracted from machinery and equipment (See main text). Also, transport equipment was not separately available for both public and private sectors since 2016, and therefore, we used a fixed proportion of transport equipment to total machinery and equipment from 2016.

¹¹ In some years transport equipment was provided as part of the machinery and equipment, categorized as 'tools, transport equipment and other fixed assets. In such cases, we use transport/tools, transport and other fixed asset ratio in the nearest year to separate transport equipment.

Total investment in each asset category is calculated as the sum of private and public sector investment in each asset. Investment in transport equipment was not available separately for the private sector industries. Therefore, we distributed the NAS total machinery in the private sector using the industry distribution of machinery and transport equipment. These estimates are then subtracted from each industry's total machinery & transport equipment data to obtain the transport equipment investment in the private sector industries. Using this approach, we could generate investment in transport equipment by industries for the period 1950-2016. We assumed the share of transport equipment in the total machinery and equipment group in 2016-2017 to remain constant to create nominal investment in transport equipment in the subsequent years. The nominal investment was deflated using GFCF deflator for transport equipment (see section (c)). Since the distinction between registered and unregistered manufacturing is not available in the NAS since 2016, we used a quick fix to split these two. NAS provides a division between the corporate sector, public sector, and household sector. In the current version of the data, we consider the household sector as a proxy for unregistered and the corporate and public sector as a proxy for the registered sector.

As mentioned earlier, the updates for years since 2017-2018 are solely based on published aggregate GFCF data from the National Accounts. In the 2019 version of the India KLEMS, we considered the 2012-2016 detailed asset-wise data by industries provided by the CSO as such and extended the data for 2017-2018 using the published aggregates. However, in the 2020 version of the data, we changed this approach, as there were some discrepancies between the published NAS totals and the detailed asset-industry data. While extending the past series of detailed asset-wise investments forward for years since 2017-2018 using the NAS-published headline series, we now consider the published NAS series as the benchmark since 2012. Then we apply the industry distribution of the 2012-2016 detailed series obtained previously from the CSO. The current approach is more appropriate and ensures a complete consistency with the NAS-published series. However, it may cause some changes in the capital service growth rates compared to the last release¹².

(b) *Asset-wise investment for non-NAS sectors*

NAS provides data only for 9 broad sectors, while we have 27 industries, which necessitated further splitting of some of the NAS sectors. This includes aggregate manufacturing (registered and unregistered separately for the period 1950-2016) with 13 sub sectors; other services into 4 sub sectors; and real estate activities and business services into 2 sub sectors. The manufacturing sector investment data was disaggregated into 13 subsectors at the 2-digit level of NIC 1998 using ASI and NSSO data, which will be discussed in detail subsequently. Investment series in service sector has been split into sub sectors using two alternative approaches—value added shares, and

¹² A detailed discussion on how the series prior to 2011-2012 were adjusted to match the subsequent series is provided in Box 5.2A in the Appendix.

capital/labour ratio in the higher aggregate industry. However, the final data used are based on value added shares, as we did not see any significant difference between the two.

Registered (organized) Manufacturing: In order to split the aggregate capital formation in organized manufacturing sector into 13 study sectors, we use the ASI. ASI defines GFCF as actual additions (newly purchased, second hand and own construction) minus deductions plus depreciation adjustment for discarded assets during the year. This approach is based on a single year's sample and helps to avoid potential huge negative investment series, and is also consistent with published ASI GFCF series.

The yearly detailed volumes beginning 1964-65 were used to derive the gross fixed capital formation by asset type directly. For the years 1964-1978, the relevant data are obtained from published detailed volumes. For the period, 1983-84 to 2004-05 ASI has generated detailed tables from Block C of ASI schedule that contain data on fixed assets. Data for missing years are interpolated using the changes in investment using book value method¹³. Table 5.2 provides an overview of the asset categories available in ASI, and the relevant asset categories in our study to which they are attributed. Though ASI provides investment in land, for reasons of NAS consistency we exclude it from our database.

Table 5. 2: Asset Types in ASI and India KLEMS

ASI Asset Types	India KLEMS Asset Types
Land	Excluded
Buildings	Buildings and Construction
Plants & Machinery	Machinery & Equipment
Transport Equipment	Transport Equipment
Computer Equipment including Software (from 1998)	Machinery & Equipment*
Pollution control equipment (from 2000)	Machinery & Equipment

Note: * Computer equipment and software are included in the machinery and equipment, as we are not distinguishing between ICT and non-ICT assets in this version.

Source: CSO, NAS Different Issues

Once investment in each of these assets and industries are generated using ASI data, we apply this industry-asset distribution to the published aggregate NAS GFCF series for organized manufacturing sector. It may also be noted that from 1960-61 to 1971-72, ASI data are for the census sector and from 1973-74 onwards they are for the factory sector. In order to make these two series comparable over years, we convert the data prior to 1972 to factory sector using the

¹³ Gross investment is estimated as the difference between book value of asset in period t and in period t-1 and add depreciation in period t to that.

factory/census ratio in 1973. Thus, after these adjustments, we obtain investment data for 13 manufacturing sectors, by asset types, consistent with the NAS aggregate for registered manufacturing. This approach was accurate for the period 1950-2016 when the NAS provided aggregate registered manufacturing GFCF data. For the years after 2016, since when NAS does not provide registered manufacturing aggregate, we proxy registered manufacturing by the sum of public sector and corporate sector.

Unregistered (unorganized) Manufacturing: The data required for creating the gross investment series for the 13 sectors of the unorganized manufacturing sector are obtained from various rounds of NSSO surveys on unorganized manufacturing. We use 6 rounds of NSSO surveys that cover the period 1989-2016. These are 45th round (1989-90), 51st round (1994-94), 56th round (2000-01), 62nd round (2005-06), 67th round (2010-11) and 73rd round (2015-16). Unit level data has been aggregated to 13 industries using the appropriate concordance. NSSO provides net addition to owned assets during the reference year within the block of fixed assets, and we use this as a measure of our investment. The investment series arrived at for six rounds were interpolated to obtain the annual time series of unorganized gross fixed capital formation by asset type. As in the case of registered sector, once the investment by asset types across industries are constructed, the asset-industry distribution is applied to the published NAS aggregate GFCF in unregistered manufacturing to obtain NAS consistent GFCF by asset type and industries.

(c) Investment Prices by Asset Types

In order to compute asset wise capital stock using PIM (equation 5.3) and rental price (see Box5.1A in the Appendix), we require asset wise investment price deflators. Since CSO has provided us with investment data by industries and assets both in current and constant prices, we could derive the price deflators with base 2011-2012. For years before 2011, prices were spliced using 2004-2005 base investment deflators. These deflators are directly used for all the three asset categories we have. Since there was no separate asset wise data available for transport equipment in 2017-2018, the investment deflator for transport equipment from 2016-2017 was extrapolated using the trend in the investment price of total machinery & equipment, obtained from NAS.

(d) Initial Stock, Depreciation Rates and Rate of Return

For the implementation of PIM to estimate asset wise capital stock, we require an estimate of initial benchmark stock (see Erumban, 2008b for an in-depth discussion on this issue). NAS provides estimates of net capital stock since 1950 for all the broad sectors in its Statement 17: Net Fixed Capital Stock by industry of use. We take the NAS estimate of real net capital stock in 1950 (in 1999-2000 prices) as our benchmark stock for all non-manufacturing sectors, and for

manufacturing sectors the same is taken for the year 1964.¹⁴ However, since the NAS estimate is available only for broad sectors and for aggregate capital, we use our industry-asset distribution of GFCF in order to create net fixed capital stock estimates by asset type for all the 27 sectors.

The approach to split asset-wise capital stock is changed in the 2020 version of the India KLEMS database. Instead of using the GFCF distribution, we use the distribution of net capital stock by assets, available from the detailed asset data obtained from the CSO. Since the initial stock depreciates over the years, the impact of this change in distribution on the capital stock will vanish over the years at the individual asset level. However, as it changes the distribution of assets in the total capital stock, these compositional changes tend to have a visible impact on the aggregate capital stock, causing a historical revision of capital stock growth rates in some industries (esp. transport and storage, education, hotels and restaurants, electricity, transport equipment, etc.).

NAS also provides detailed tables on assumed life of assets used for computing capital stock, for private units, administrative units as well as departmental and non-departmental units by asset types.¹⁵ We use these estimates of lifetime to derive appropriate depreciation rates for non-ICT assets, using a double declining balance rate. Following the NAS, we assume 80 years of lifetime for buildings, 20 years for transport equipment, and 25 years for machinery and equipment (see Appendix 26.2; CSO, 2007). The final depreciation rates used in the study are given in Table 5.3 by asset type. Subsequently, we build our capital stock series by asset types for all the 27 industries using our GFCF series from 1950 (1964) onwards for the non-manufacturing (manufacturing) sectors.

Table 5.3: Depreciation Rate by Asset Type Used in the Computation of Capital Input

Asset Type	Depreciation Rate (%)
Building and Construction	2.5
Transport Equipment	10.0
Machinery	8.0

Note: Depreciation rates are derived using NAS life times for each asset assuming a double declining balance rate.

Our measure of capital input is arrived using equation (5.1), for which we also require estimates of rental prices (see Box 5.1A in the Appendix). Assuming that the flow of capital services is proportional to the capital stock at individual asset level, aggregate capital flows can be obtained using a translog quantity index by weighting growth in the stock of each asset by the average shares of each asset in the value of capital compensation, as in (5.1). The rate of return (i) in equation (5.4) represents the opportunity cost of capital, and can be measured either as internal (or

¹⁴ This choice is driven by the fact that the first year of availability of ASI data is 1964-65.

¹⁵ National Accounts Statistics-Sources and Methods, Chapter 26, CSO (2007),

ex post) rate of return, or as an external (*ex ante*) rate of return.¹⁶ The present version of the database uses an external rate of return, proxied by average of return on government securities and prime lending rate obtained from the Reserve Bank of India¹⁷. Therefore, we use a real rate, which is net of capital gain. Hence, the capital gain component in equation is excluded while estimating rental price using external rate of return, obtaining

$$P_{k,t}^K = P_{k,t-1}^I \cdot r_t + P_{k,t}^I \cdot \delta_k \quad (5.4)$$

Where r is the real rate of return, nominal interest rate adjusted for consumer price inflation rate. The consumer price indices (CPI) are obtained from IMF and World Bank.

5.3 Outstanding issues

The measures of capital input available in India KLEMS are based on only three asset categories, construction, machinery, and transport equipment. Therefore, it does not consider the enormous and distinct role of ICT capital in contributing to productivity and growth. The absence of ICT estimates in the database is driven by insufficient data on investment in ICT assets, such as hardware, software, and communication equipment in the National Accounts. However, during recent years, NAS has extended its asset coverage to include software and ICT equipment. Erumban and Das (2016) have made some estimates for the aggregate economy using input-output tables for ICT equipment and relying on NAS for software investment. Similarly, Erumban and Das (2020) have made some initial estimates of ICT capital by industries, which still remains a challenging task. Nevertheless, now that more data is available from NAS, it is possible to explore this further, and future extensions of the database may explore this. Similarly, following the SNA 2008 guidelines, the NAS now provides estimates of intangible capital such as intellectual property and research and development (R&D), which may be a useful extension to the capital input database.

Two additional challenges are arising since the new system of national accounts. The first is the lack of separate data on registered and unregistered manufacturing, which poses challenges in using ASI and NSSO data to split aggregate manufacturing data into 13 industry groups. Instead of the registered and unregistered split, CSO now provides a division between the household and corporate sectors. We currently address this challenge by considering the household sector as unregistered manufacturing and the sum of public and corporate sectors as a registered sector.

The second challenge is the lack of separate data on transport equipment and machinery assets in the new series of national accounts, which follows United Nation's System of National Accounts (SNA)-2008. Although, for the time being, we adopt ad-hoc approaches to fix this, it cannot be

¹⁶ We do not intend to delve into the controversies over the use of internal vs. external rate of return in the context of productivity measurement. This database uses See Erumban (2008a and b) for a discussion on these issues.

¹⁷ Reserve Bank of India, Handbook of Indian Statistics, Annual volumes.

adopted on a long-term basis, which might necessitate combining transport equipment and machinery to one single asset.

Finally, as mentioned earlier, the debate on whether it is appropriate to use an external rate of return or an internal rate of return is unsettled in the literature. It may be worth exploring estimates of capital using the internal rate of return.

Appendix B: Some Discussions on the Estimation of Capital Services

Box 5.1A: Estimation of Rental Price of Capital

The rental price of capital $P_{k,t}^K$ reflects the price at which the investor is indifferent between buying and renting the capital good for a one-year lease in the rental market. In the absence of taxation, the rental price equation can be derived as the sum of the nominal rate of return, the nominal cost of depreciation, less capital gain (see Jorgenson and Griliches, 1967; and Christensen and Jorgenson, 1969):

$$P_{k,t}^K = P_{k,t-1}^I \cdot i_t + P_{k,t}^I \cdot \delta_k - (P_{k,t}^I - P_{k,t-1}^I) \quad (5.4)$$

where $P_{k,t}^I$ is the investment price (or acquisition price) of asset k in year t , and i_t is the nominal rate of return. The first component is the rate of return component, which captures the opportunity cost of purchasing a capital asset. The second component is the depreciation component, and the last component, which is the difference between investment prices in year t and $t-1$, captures the capital gain. Thus, this formula shows that the rental price is determined by the nominal rate of return, economic depreciation costs, and asset-specific capital gains.¹⁸ The rate of return i in equation (5.4) can be measured either as an internal or ex-post rate of return – using the information on realized capital compensation – or an external or ex-ante rate of return – such as the prevailing long-term interest rate in the economy. There is continuing debate on which rate is more appropriate for productivity analysis. Empirical evidence suggests that the choice of a particular rate may impact final estimates of capital services and productivity (see Oulton, 2007; Erumban, 2008a). In India KLEMS, following the arguments that investors make their investment decisions considering the prevailing interest rates in the economy, we use an external rate of return (see more details in section 2.4). Ideally taxes should be included to account for differences in tax treatment of the different asset types and different legal forms (household, corporate and non-corporate). The capital service price formulas above should then be adjusted to take these tax rates into account. However, this refinement would require data on capital tax allowances and rates by industry and year, which is beyond the scope of this database. Available evidence for major European countries shows that the inclusion of tax rates has only a very minor effect on growth rates of capital services and TFPG (Erumban, 2008a).

¹⁸ The economic rationale of using the rental prices to calculate a reliable service growth is that the investor expects to get more services in short time from an asset whose price is relatively high (or service life is relatively small). In equilibrium, an investor is indifferent between two alternatives: earning a nominal rate of return i on an investment or buying a unit of capital collecting a rental P^K and then selling it at the depreciated asset price $(1-\delta)P^I$ in the next period. Assuming no taxation, the equilibrium condition is: $(1+i_t)P_{k,t-1}^I = P_{k,t}^K + (1-\delta_k)P_{k,t}^I$, with P^K as the rental fee and P^I the acquisition price of investment good k (Jorgenson and Stiroh 2000, p.192). Rearranging yields the cost-of-capital equation: $P_{k,t}^K = P_{k,t-1}^I \cdot i_t + \delta_k P_{k,t}^I - [P_{k,t}^I - P_{k,t-1}^I]$.

Box 5.2A: Backcasting of NAS Data-Methodologies

As we use multiple vintages of GFCF data from various revisions of national accounts, we extrapolate the 2011-2012 base nominal GFCF series backwards using growth rates from the 2004-2005 series. Using. While doing this, in the previous versions of the KLEMS data, we first extrapolated each individual industry in the NAS detailed data and then aggregated them to the KLEMS sector. For instance, suppose we have n NAS industries ($i=1, \dots, n$) that constitute a KLEMS sector j (e.g., the KLEMS sector transport & storage consists of NAS sectors railways, air transport, water transport, road transport, other public transport, other private transport, other transport services, and storage). In the previous versions of the data, we extrapolated the nominal GFCF in the 2012 series for each of these n sectors individually backward using the growth rates from the nominal GFCF series from the 2004 version. We change this approach in the current 2020 version to extrapolate after aggregating detailed data to the KLEMS sector, as we observe significant inconsistencies in the detailed industry data for some industries between the two versions of the NAS. Therefore, we first aggregate the nominal GFCF for the KLEMS sector and then extrapolate it backward for each KLEMS industry. More formally, the approach we followed in KLEMS version 2019 was:

$$GFCF_{j,t}^{2012} = \sum_{i=1}^n GFCF_{i,t}^{2012}; \text{ for all years since 1950}$$

where $GFCF_{j,t}^{2012}$ is the nominal GFCF in KLEMS industry j , and $GFCF_{i,t}^{2012}$ is the GFCF in NAS industry i , both for year t and in 2012 base. The former is obtained as:

$$GFCF_{i,t}^{2012} = \left[\frac{GFCF_{i,t}^{2004}}{GFCF_{i,t+1}^{2004}} \right] GFCF_{i,t+1}^{2012} \text{ for years 1950-2011, and they are directly obtained from the NAS for years after 2011.}$$

In the current 2020 version of the KLEMS data, the approach to obtain GFCF for the KLEMS industry j is:

$$GFCF_{j,t}^{2012} = \sum_{i=1}^n GFCF_{i,t}^{2012} \text{ for years 2016-2011, and}$$

$$GFCF_{j,t}^{2012} = \left[\frac{GFCF_{j,t}^{2004}}{GFCF_{j,t+1}^{2004}} \right] GFCF_{j,t+1}^{2012} \text{ for years 1950-2011}$$

$$\text{where } GFCF_{j,t}^{2004} = \sum_{i=1}^n GFCF_{i,t}^{2004}$$

The change in the aggregation procedure causes some differences in the historic GFCF growth rate for the KLEMS sector j for the period 1950-2011. This difference in the GFCF growth also leads to a historical revision of capital stock growth rates for the entire series in some industries (e.g., agriculture, which is the sum of crops, fisheries, forestry; utility sector, which is the sum of electricity, gas, and water supply; transport services, which is the sum of different transport

services, storage, etc.), as capital stock is a cumulative sum of investment stream over the years. While it may be accurate to use detailed industry data to extrapolate backward (the earlier approach), the magnitude of the fluctuation might affect the quality of the data. Moreover, the current approach is consistent with the approach taken for gross value added in the KLEMS database.

Chapter 6: Intermediate Input Series at the Industry Level

In this section, we describe the basic approach we have used to derive the volume series of Intermediate Inputs namely –Energy input (E), Material input (M) and Services input (S).

6.1 Methodology

The methodology for measuring industry output, intermediate inputs and value added was developed by Jorgenson, Gallop and Fraumeni (1987) and extended by Jorgenson (1990 a). The cornerstone of this approach is a time series of input output (IO) tables which gives the flows of all commodities in the economy, as well as payments to primary factors. Following a similar approach as explained in Jorgenson *et al.* (2005, Chapter 4) and Timmer *et al.* (2010, Chapter 3), the time series on intermediate inputs for the India KLEMS project have been constructed.

Definition of EMS: As in EU KLEMS, this study identifies three main categories of Intermediate inputs. They are classified as follows:

1. **Energy Input**
2. **Material Input**
3. **Service Input**

The following five energy types (and products) have been classified as the **Energy input**:

1. Coal and Lignite
2. Petroleum products
3. Electricity; (for electricity used in the electricity sector, since there is a good amount of inter-firm sale and purchase of electricity, it has been treated as material rather than as energy)
4. Natural gas
5. Gas (LPG)

The following fourteen input items have been classified as the **Service input**:

1. Water supply
2. Railway transport services
3. Other transport services
4. Storage and warehousing
5. Communication
6. Trade
7. Hotels and restaurants

8. Banking
9. Insurance
10. Ownership of dwellings
11. Education and research
12. Medical and health
13. Other services
14. Public administration

All other intermediate inputs barring the above mentioned nineteen inputs are classified as **material input**.

6.2 Implementation Procedure

The methodology for computation of Intermediate Input Series for 27 Industries from 1980-2019 at current and constant prices is explained in steps.

Step 1: Concordance is done between IOTT and study industries

- The 60/115/130 IOTT industries are aggregated to form industries.
- A concordance table between the classification used in our study and each Input Output Transaction Table has been prepared. (refer KLEMS Data Manual Sept 2021 release for concordance table of KLEMS industries with IOTT 1998 and 2003)

Step 2: Obtaining estimates for Material, Energy and Service Inputs for 27 Industries, in benchmark years

1. In IOTT, no intermediate input is being used in industry 24 – ‘Public Administration’. Consequently, the intermediate input series is being directly estimated for only 26 industries from 1980 to 2019. However, we have used ‘Net Purchase of Commodities and Services’ of Administrative Department as total intermediate input. And using information about ‘Government Final Consumption Expenditure’ we distribute the total intermediate input in to material, energy and services input.
2. Value of Energy Inputs used
3. Value of Material Inputs used
4. Value of Service Input used
5. Value of Total Intermediate Inputs (summation of the above three)

Thus, for each of the benchmark year, estimates are obtained for Material, Energy and Service Inputs that has been used to produce Gross output in the 27 different India KLEMS Industries.

Step 3: Projecting a time series (1980 to 2019) of proportions of Material, Energy and Service Inputs in Total Intermediate Inputs for each of the 27 industries

- For the Benchmark years *i.e.*, 1978, 1983, 1989, 1993, 1998, 2003, 2007 and 2015 proportions of Material Inputs, Energy Inputs, and Service Inputs in Total Intermediate Inputs are calculated.
- Similar proportion for intervening years is obtained by linear interpolation of the benchmark proportions.
- This gives a time series of proportions of Material, Energy and Service Inputs in Total Intermediate Inputs for each year, from 1980 to 2019.
- Thus, using IOTT, the intermediate input vector has been projected for 27 study industries from 1980 to 2019.

There exists some abnormal fluctuation in the proportions of Material Inputs, Energy Inputs, and Service Inputs in Total Intermediate Inputs for benchmark years. In that case, we drop that specific benchmark proportion and linearly interpolate of the adjacent benchmark proportions to estimate proportion for intervening years. For the industry ‘Wood & Wood Products’, we estimated proportions of Material Inputs, Energy Inputs, and Service Inputs in Total Intermediate Inputs from SUT 2015-16 instead of Input Output Transaction Table 2015-16.

For the industry group ‘Coke, Refined Petroleum and Nuclear Fuel’, we estimated proportions of Material Inputs, Energy Inputs, and Service Inputs in Total Intermediate Inputs from ASI data instead of Input Output Transaction Table. This has been done because the relevant proportions differed significantly between the input-out tables and ASI, and the latter is believed to be making a more correct assessment of energy inputs used in the Coke and Petroleum products industry.

Step 4: Consistency with NAS

The projection of intermediate input vector, using IOTT in Step 3 needs to be consistent with the estimated output from NAS.

- The Gross Value of Output and the Gross Value Added at current prices calculated from NAS is taken. The difference between GVO and GVA provides us the gap between value added and gross output for each year or in other words this reflects the Total Intermediate inputs from NAS (GVO-GVA).
- The intermediate input vector that has been projected using IOTT in the step 3 will not tally with the above estimates of Total Intermediate inputs from NAS. Thus, the projected input vector, has been proportionately adjusted to match the gap between value added and

gross output obtained from NAS. This involves two steps:

1. **For Benchmark years:** Ratio of Total Intermediate inputs from NAS to that from IOTT is adjusted proportionately to the absolute value of Energy Inputs/Material Inputs/Service Inputs obtained from IOTT.
2. **For Intervening Years:** The interpolated proportions of Energy Inputs, Material Inputs and Services Inputs obtained from IOTT, is applied directly to the total intermediate inputs from NAS to get each inputs share.

Steps 1-4: This gives a time series of Material, Energy, and Service Inputs for 27 study Industries from 1980 to 2019 at current prices.

Steps 5 and 6 below, explain the methodology for computation of Intermediate Input Series for 27 study Industries from 1980 to 2019 at constant price. The approach followed here is to first form the aggregates of materials, energy and services at current price for each study industry from the benchmark Input Output tables and then develop deflators of Materials, Energy and Service Inputs for each of the 27 study Industries separately.

Step 5: Constructing Deflators of Materials, Energy and Service Inputs for 27 study Industries separately

- Deflators are obtained for each of the 115 (IOTT 1989 and 1998) and 130 (IOTT 2007) commodity inputs (each row of the IO Matrix).
- WPI for the period 1980 to 2019 is taken from Office of the Economic Adviser, Ministry of Commerce and Industry.
- Best available Wholesale Price Index is applied to each commodity input in the IOTT.
- In several cases, no proper WPI is available for the sector; hence the best among the available has been chosen.
- In some cases, the right index has been formed from item level indices.
- In some case, one index has been removed from a higher-level aggregate to generate the right index.
- For Electricity as an input entering into the production process of an industry, depending on the nature of economic activity, the right price of electricity has been chosen. For instance, in the case of manufacturing industries, the price of electricity for industry considered, similarly for agriculture the price of electricity for agriculture.
- For natural gas from 2004 to 2020 the price information is taken from Indian Petroleum and Natural Gas Statistics. The detailed method of price index construction for Natural gas for the period 1980-81 to 2019-20 is given in box 1 below :

Box 1: Natural gas price index

While data on prices of natural gas are available for the recent period (for the period from 2004-05 onwards, regular data exist on natural gas prices), very little data exists on natural gas prices for the earlier period, particularly for the 1980s. Therefore, the scanty information that is available had to be used, and additionally, assumptions and interpolations had to be resorted to for constructing the price index series. The methodology that has been adopted is discussed below. The price series is constructed for the period 1980-81 to 2019-20. This period is divided into (a) 2003-04 to 2019-20, (b) 1992-93 to 2003-04, (c) 1987-88 to 1992-93 and (d) 1980-81 to 1987-88.

A. Period: 2003-04 to 2019-20

Data on consumer price for natural gas (in Rs/ '000' cubic meters) for different years from 2004 to 2020 are available in the *Indian Petroleum and Natural Gas Statistics*, Ministry of Petroleum and Natural Gas, Government of India. For 2003-04, the price information is available for January 2004 and March 2004. This price has been taken as the price for the year 2003-04. The ratio between Off-shore price at landfall point (applicable to on-shore) and the price for north-eastern states maintains a steady ratio (of about 60%) in the period 2003-04 to 2019-20. Therefore, the Off-shore price is used to prepare the price index. The base year of the price index is taken as 2011-12=100.

B. Period: 1992-93 to 2003-04

Price for natural gas for the four years 1992-93, 1993-94, 1994-95 and 1995-96 have been taken from Negi, Pahwa and Arora (2012). A weighted average price for the years 1992-93 to 1995-96 have been taken combining the All-India price and the price for north-eastern states. T.L. Shankar committee report in 1996 recommended that price of natural gas should be linked to international fuel oil basket. The linkage was specified as 65% for 1998-99, 75% for 1999-2000 and 2000-01 and 100% 2001 onwards. However, there was a rider that the price of natural gas is to be within a band of Rs. 2150/Kscm to 2850/Kscm for gas with CV=10,000 Kcal/scm and for north-eastern states it will be Rs. 1200- 1700/Kscm. In July 2003, the ceiling price was raised to Rs. 3200/Kscm for Fertilizer and Rs 3600/Kscm for others, although at 100% linkage (as proposed by TL Sankar Committee) it would have been Rs 8140/Kscm (source: Negi, et al. 2012). However, according to data provided in *Indian Petroleum and Natural Gas Statistics*, the price remained at Rs 2850 and Rs 1700 for All-India and for North-eastern state, respectively, in 2003-04 and 2004-05. It is only in the middle of 2005-06 that the price for All-India was raised to Rs 3200/Kscm.

Considering the above information on the price on natural gas, the required data on natural gas prices are available for 1995-96 and 2003-04. For the intervening years, 1998-99 to 2002-03, the price has been taken within the range mentioned above with the help of certain assumptions. The assumption is about which point of the specified range (Rs 2150-2850 per Kscm) the price might have been fixed taking into account the link with international fuel oil prices. For 1998-99, the bottom of the range is taken. For 1999-2000 and 2001-02, the mid-point of the range is taken, and for 2002-03, the end point of the range is taken. For 2003-04, the price has been fixed on basis of the price data for January 2004 and March 2004 when it was at the top end of the range. Since the required information is not readily available, this is the best that could be done.

The T.L.Shankar Committee Report is dated December 1996 and the recommendations were effective from 1-10-1997. Thus, the price range given by the committee could not be applied to 1996-97 and 1997-98. For these two years, the price has been taken to be same as in 1995-96.

C. Period: 1987-88 to 1992-93

According to the Report of the Shankar committee, it was decided in 1986 that the price of natural gas will be henceforth fixed by the government. Thus, price information is available for 1987-88 and 1992-93. The prices for these two points have been interpolated to obtain the price for intervening years. Interpolation has been done considering the movements in the wholesale price index for crude oil.

D. Period: 1980-81 to 1987-88

For this period, there is very little information on natural gas price. It is known that there was a wide range in the price charged. From the published reports of ASI, the unit value of purchase of natural gas in fertilizer industries in 1983-84 is obtained. Using ASI data for 1983-84 and the wholesale price index for crude oil and natural gas, an index for natural gas price has formed for this period, which has then been spliced with the price indices of the three periods mentioned above.

E. Final series on price index

Having obtained the price index for the four sub-periods as explained above, these have been spliced to obtain the price index for natural gas for the period 1980-81 to 2019-20.

- For Service Inputs, since WPI is not available hence implicit GDP deflators from NAS are used.
- Deflators obtained for different IO sectors have been combined using weights. The weights are based on the column of the relevant study industry in the IO table. This is because; the entries read down industry columns give the commodity inputs of raw-materials and services, which are used to produce outputs of particular industries. We have allowed for changing weights over time *i.e.*, we have used different weights for different time periods. Three IOTT has been used for this purpose – 1989, 1998 and 2007. The price series based on 1989 table has been used from 1980 to 1993, the 1998 table has been used from 1993 to 2003 and the 2007 table has been used for the price series for the period 2003 to 2019. Once the three series have been formed, these have been spliced. Thus, the IO tables have been used for obtaining the materials, energy and services series for each industry at current prices

Step 6: Computing Time Series on Intermediate Input for 27 study Industries from 1980-2019, Constant prices

- The deflators for Material, Energy and Service Inputs for each study industry have been used to deflate the Current price Intermediate Input series to Constant price.
- Gross Value Added for Agriculture estimated by NAS using double deflated method. Therefore, for the Agriculture sector the deflated value of intermediate input is adjusted to the gap between value added and gross output at constant price.

Thus, we have a time series of Material, Energy and Service Inputs for 27 study Industries at constant prices.

Chapter 7: Factor Income Share Series at the Industry Level

The distribution of income between capital, labour and intermediate inputs, is an important element in growth accounting because income shares, under conditions of competitive markets, can be used to measure the contributions each factor makes towards output growth.

7.1 Methodology for Measuring Labour Income Share Series

Under the assumption of constant returns to scale with two factors of production *i.e.*, labour and capital, the sum of the labour income share and capital income share is 1. The labour income share is defined as the ratio of labour income to GVA. Capital income share is accordingly obtained as one minus labour income share.

There are no published data on factor income shares in Indian economy at a detailed disaggregate level. *National Accounts Statistics* (NAS) of the CSO publishes the NDP series comprising of compensation of employees (CE), operating surplus (OS) and mixed income (MI) for the NAS industries. The income of the self-employed persons, *i.e.*, mixed income (MI) is not separated into the labour component and capital component of the income. Therefore, to compute the labour income share out of value added, one has to take the sum of the compensation of employees and that part of the mixed income which are wages for labour.

The computation of labour income share for the 27 study industries involves two steps. First, estimates of CE and MI have to be obtained for each of the 27 study industries from the NAS data which are available only for the NAS sectors (see Table 7.1). Second, the estimate of mixed income has to be split into labour income and capital income for each industry for each year (except for those industries for which the reported mixed income is zero, for instance, public administration).

Table 7. 1: NAS Sectors and Corresponding Study Industries, for Computation of Labour Income Shares

NAS sectors for which factor income data are available		Corresponding Study Industries
1. Agriculture, forestry & logging and fishing	1	Agriculture, Hunting, Forestry and Fishing
2. Mining & Quarrying	2	Mining and Quarrying
3.1 Manufacturing-registered	3	Food Products, Beverages and Tobacco
3.2 Manufacturing - unregistered	4	Textiles, Textile Products, Leather and Footwear
Both registered and unregistered manufacturing are split into 13 manufacturing industries (3 to 15)	5	Wood and Products of Wood
	6	Pulp, Paper, Paper Products, Printing and Publishing
	7	Coke, Refined Petroleum Products and Nuclear Fuel
	8	Chemicals and Chemical Products
	9	Rubber and Plastic Products
	10	Other Non-Metallic Mineral Products
	11	Basic Metals and Fabricated Metal Products
	12	Machinery, nec.
	13	Electrical and Optical Equipment
	14	Transport Equipment
	15	Manufacturing, nec; recycling
4. Electricity, Gas and Water Supply	16	Electricity, Gas and Water Supply
5. Construction	17	Construction
6.1 Trade	18	Trade
6.2 Hotels and Restaurants	19	Hotels and Restaurants
7.1 Railways	20	Transport and Storage
7.2 Transport by other means		
7.3 Storage		
7.4 Communication	21	Post and Telecommunication
8.1 Banking and Insurance	22	Financial Services
9.1 Public administration and Defense	24	Public Administration and Defence; Compulsory Social Security
9.2 +8.2 Other Services plus	25	Education
Real Estates, Ownership of Dwelling and Business Services	26	Health and Social Work
	23	Business Services

*Note: *Study Industry 'Other Services' includes Real Estate Activities; Other Community, Social and Personal Services, Private Household with Employed Persons.*

Basis data sources used for the computation of labour income share are NAS, ASI and unit level data of survey of unorganized manufacturing enterprises. These data sources are used to obtain estimates of CE, OS and MI for each of the 27 study industries. For splitting the labour and non-labour components out of the mixed income of self-employed, the unit level data of NSS employment-unemployment survey are used along with the estimates of CE, OS and MI basically obtained from the NAS.

7.2 Implementation Procedure

a) Construction of Labour Income Share Series in Gross Value Added

The estimation of labour income share for the 27 study industries has been done in two steps, as discussed below.

Step 1: Estimation of CE, OS and MI for the 27 study Industries

NAS provides estimates of compensation of employees (CE), operating surplus (OS) and mixed income (MI) for the 9 NAS industries in different annual Sequence of National Accounts reports.

For some industries under study, for instance (i) Agriculture, forestry & logging and fishing, (ii) Mining & Quarrying, (iii) Electricity, gas and water supply and (iv) Construction, the required data are readily available from NAS. For others, the estimates available in NAS have to be distributed across the study industries using ASI data. The NAS estimates of factor incomes (CE) for registered manufacturing is split into KLMES 13 industries (3-15) in proportion to the reported ASI data on emoluments for various industries. Estimate of factor incomes for 'other services' in NAS has been split into estimates for (i) Education, (ii) Health and Social Work, and (iii) Other Community, Social and Personal Services including Renting of Machinery and Business Services, and Private Household with Employed Persons.

Before 2011-12, NAS provides estimates of factor incomes for registered manufacturing and unregistered manufacturing, but not for individual manufacturing industries. However, 2011-12 onwards NAS disaggregated the manufacturing sector in corporate sector and household sector. The NAS estimates of factor incomes for registered manufacturing and corporate sector have to be split into various manufacturing industries considered in the study (13 in number) using ASI data. The reported CE in NAS for registered manufacturing and corporate sector has been

distributed into those 13 industries in proportion to the reported ASI data on emoluments for various industries. In a similar way, using ASI data, the estimate of OS for registered manufacturing and corporate sector has been distributed. Emoluments are subtracted from gross value added for various industries yielding capital income. The share of different industries in aggregate capital income of organized manufacturing indicated by the ASI data is used to split the estimate of OS for registered manufacturing and corporate sector reported in the NAS.

The methodology applied for unregistered manufacturing and household sector is similar. The published results and unit level data of survey of unorganized manufacturing industries have been used for this purpose. The estimates of wage payments (to hired workers) in different industries have been used to split (proportionately) the estimate of CE in the NAS for aggregate unregistered manufacturing. The estimated wage payment is subtracted from the estimated value added to obtain an estimate of capital income and mixed income of the self-employed in various unorganized manufacturing industries. The estimate of MI provided in NAS for unregistered manufacturing and household sector has then been proportionately distributed across industries using the estimate of capital income and mixed income of the self-employed in various industries that could be formed on the basis of published results and unit level data of survey of unorganized manufacturing industries.

Unlike the ASI data for organized manufacturing, the data for unorganized manufacturing enterprises are available for only select years. The proportions mentioned above could therefore be computed only for those select years (data for six rounds have been used; these are for 45th Round (1989-90), 51st Round (1994-95), 56th Round (2000-01), 62nd Round (2005-06), 67th round (2010-11) and 73rd Round (2015-16)). It has accordingly, been necessary to resort to interpolation/ extrapolation to obtain the relevant proportions for other years.

Step 2: Splitting of MI into Labour Income and Capital Income

As explained above, the income share of labour is computed as:

$$SL_t^i = \frac{CE_t^i + (\eta^i \times MI_t^i)}{GVA_t^i} \dots \quad (7.1)$$

In this equation, SL_t^i is the labour income share in industry i in year t , CE_t^i is compensation of employees in industry i in year t , MI_t^i is mixed income of the self-employed persons in industry i in year t , and GVA_t^i is gross value added in industry i in year t . The labour income proportion in income is denoted by η which taken to be a fixed parameter for each industry, not varying over time.

The derivation of the GVA series for different industries has been briefly explained in Chapter 2. The derivation of CE and MI series has been explained in step I above. Therefore, only the estimation method of η needs to be described. The estimation of η has been done with the help of NSS survey-based estimates of employment of different categories of workers (number of persons and days of work) and wage rates (which has been described briefly in chapter 4) coupled with estimates of MI, basically obtained from the NAS. Two approaches have been taken to get an estimate of η , and the labour income share series for different industries finally adopted in the study makes use of an average of the estimates of η obtained by the two approaches.

In the first approach, an estimate of labour income of self-employed workers has been made for each study industry for eight years, 1983-84, 1987-88, 1993-94, 1999-00, 2004-05, 2009-10, 2011-12 and 2015-16 on the basis of the estimated number of self-employed, wage rate of self-employed and the number of days of work per week. The industry-wise estimates of number of self-employed, wage rate of self-employed and the number of days of work per week have been made from unit records of NSS employment-unemployment survey (major rounds) of 1983, 1993-94, 1999-00, 2004-05, 2011-12, and PLFS of 2018-19 and 2019-20.

The estimates of the number of self-employed, wage rate of self-employed and the number of days of work per week provide an estimate of the annual labour income of self-employed workers which is divided by the mixed income of self-employed (derived from NAS) to get an estimate of η . For five industries, the ratio in question has been computed and applied. For the other 22 industries, the ratio in question has been computed after clubbing the industries into 11 industry groups.¹⁹ In the latter case, a common ratio computed for group of industries has then been applied to constituent industries. The list of industries or industry groups for which η has been estimated is given in table 7.2.

In the second approach, the NSS data are used to compute the following ratio: the ratio of labour income of self-employed workers to the labour income of regular and casual workers. Let this be denoted by θ . Then, the estimate of CE provided in the NAS is multiplied by θ to obtain an estimate of the labour income component out of the MI reported in the NAS. The labour component of MI divided by total MI gives an estimate of η . In case the estimated labour component of MI exceeds the estimate of MI, the estimate of η has been taken as unity.

Examining the estimates obtained by the first approach, it was found that the estimated labour income share out of mixed income varied significantly among the estimates for the seven years for which the ratio in question has been estimated.

¹⁹ The estimation is done at group level rather than individual industries on the consideration that the group level estimates will be more reliable.

The estimates of η obtained by the second approach has the problem that in a number of cases, the estimated labour component of MI exceeds the estimate of MI given in the NAS, and therefore η is taken as one.

The method finally adopted is as follows: (a) The average value of η has been computed for each industry or industry group by taking the estimates for the years 1983-84, 1987-88, 1993-94, 1999-00, 2004-05, 2009-10, 2011-12 and 2015-16. This has been done separately for the estimates based on approach-1 and those based on approach-2. (b) The estimates of η obtained for each industry or industry group by the two approaches have then been averaged. (c) Having obtained an estimate of η , equation 7.1 given above has been applied to compute the labour income share.

Table 7. 2: Industries and Groups for which η (proportion of labour income out of mixed income) has been estimated

Industry/Industry Group	Study Industries
Agriculture, Hunting, Forestry and Fishing	Agriculture, Hunting, Forestry and Fishing
Mining and Quarrying	Mining and Quarrying
Food Products, Beverages, Tobacco, Textiles, Leather Products	Food Products, Beverages and Tobacco Textiles, Textile Products, Leather and Footwear
Wood, Wood Products, Paper and Printing	Wood and Products of Wood Pulp, Paper, Paper Products, Printing and Publishing
Petroleum Products, Chemical Products, Rubber and Plastic Products	Coke, Refined Petroleum Products and Nuclear Fuel Chemicals and Chemical Products Rubber and Plastic Products
Other Non-Metallic Mineral Products	Other Non-Metallic Mineral Products
Metals, Metal Products, Machinery and Transport Equipment	Basic Metals and Fabricated Metal Products Machinery, nec. Electrical and Optical Equipment Transport Equipment
Manufacturing, nec; Recycling	Manufacturing, nec; recycling
Electricity, Gas and Water Supply	Electricity, Gas and Water Supply
Construction	Construction
Trade, Hotels, Restaurants	Trade Hotels and Restaurants
Transport, Storage and Communication	Transport and Storage Post and Telecommunication
Public Administration & Defence	Public Administration and Defence; Compulsory Social Security
Education and Health	Education Health and Social Work
Other Services including Financial services, Real Estates, Business Services	Financial Services Business Services Other Services

Source: National Account Statistics (NAS) reports.

b) Construction of Factor Income Share Series in Gross output

The income share of labour, capital and intermediate inputs in Gross output has been computed using the following steps:

- First the individual shares of intermediate inputs that is energy, material and services in Gross output is calculated.
- The labour income and capital income out of Gross Value Added are further distributed into income share of labour in gross output and income share of capital in gross output.
- Thus, we get the share of labour, capital, material, energy and services in output.
- For Agriculture Sector, land is taken as an input. First income share of intermediate inputs in gross output is calculated. Then the non labour income (out of value added) in Agriculture is distributed equally between land and capital inputs to derive weights.

7.3. Outstanding Issues

The splitting of unorganized sector factor incomes into individual industries has been done using unorganized survey results. The proportions computed for 1983-84 has been applied for the period 1980-81 to 1983-84. While there were surveys of unorganized manufacturing enterprises in 1978-79 the survey results have not been used for estimation of factor incomes in different industries of the unorganized manufacturing sector.

Chapter 8: Growth Accounting Methodology

This chapter deals with the methodology of measurement of total factor productivity (TFP) growth for individual industries in the KLEMS framework and the aggregation from industry level productivity measures to measures for broad sectors and the economy as a whole. The methodology of analysis of sources of gross output growth at the individual industry level and sources of GVA (Gross Value Added) or GDP growth at the broad sector level and economy level will also be presented in this chapter

8.1 Methodology for Measuring Productivity Growth at the Industry level

The production function

Our measurement of TFP growth for different industries of the Indian economy is based on a gross output production function for each industry j :

$$Y = f(L(L_1, \dots, L_n), K(K_1, \dots, K_m), E(E_1, \dots, E_s), M(M_1 \dots M_u), S(S_1 \dots S_v), T) \quad (8.1)$$

Y is industry gross output, L is labour input, K is capital input and E , M and S are intermediate inputs-namely energy, material and services and T is an indicator of technology for industry j . All variables are indexed by time (t subscript is suppressed).

There are several things to note about the production function - (1) all variables are aggregates of many components that have been discussed in the earlier report/chapters; (2) we assume that the industry production function is separable in these aggregates; (3) time (indicator of technology) enters the production function symmetrically and directly with inputs. An important feature of the gross output approach is the explicit role of intermediate inputs. In our study, we have considered three intermediate inputs - energy, material and services and this is important as we may find that intermediate inputs are the primary component of some industries outputs²⁰. Failure to quantify intermediate inputs leads us to miss both the role of key industries that produce intermediate inputs and the importance of intermediate inputs for the industries that use them.

To estimate the contributions of total factor productivity, and inputs to industry level output growth, based on the above-mentioned production function, we construct real values of inputs

²⁰ Consider, the semi-conductor (SC) industry, which is a key input to the computer hardware industry. Much of the output is invisible at the aggregate level because semi-conductor products are intermediate inputs to other industries rather than deliverables to final demand - consumption and investment goods. Moreover, SC plays a role in the improvements in quality and performance of other products like-computers, communication equipment and scientific instruments. Failure to account for them leads us to miss the role of key industries that produce intermediate inputs and importance of intermediate inputs for the industries that use them. (Jorgenson, Ho and Stiroh (2005), Productivity, volume 3).

and output, using industry level price data for capital, labour and intermediate inputs, and output. e.g., P_{Y_j} ; P_{L_j} , P_{K_j} ; P_{E_j} , P_{M_j} , P_{S_j} . The industry price of an input varies across industries due to compositional effects as industries expend different shares of their investment on each type of asset. The same is true for all inputs.

Total Factor Productivity

To estimate TFP growth, we begin with the fundamental accounting identity for each industry where the value of output equals the value of inputs.

$$P_{Y_j}Y_j = P_{K_j}K_j + P_{L_j}L_j + P_{E_j}E_j + P_{M_j}M_j + P_{S_j}S_j \quad (8.2)$$

Under specific assumptions of constant returns to scale and competitive markets, we can define TFP growth as

$$v_{T,j} \equiv \Delta \ln Y_j - \bar{v}_{K,j} \Delta \ln K_j - \bar{v}_{L,j} \Delta \ln L_j - \bar{v}_{E,j} \Delta \ln E_j - \bar{v}_{M,j} \Delta \ln M_j - \bar{v}_{S,j} \Delta \ln S_j \quad (8.3)$$

Where \bar{v} is the two-period average share of the input in the nominal value of output.

We define the value share of each input as:

$$v_{i,j} = \frac{P_{X,i,j}X_{i,j}}{P_{Y_j}Y_j}; \quad i = K, L, E, M, S \text{ is the value share of input in the nominal output.} \quad (8.4)$$

and $X_{K,j} = K_j$, $X_{L,j} = L_j$, $X_{E,j} = E_j$, $X_{M,j} = M_j$, $X_{S,j} = S_j$;

$$\text{and } \bar{v}_{i,j} = 0.5 (v_{i,j,t} + v_{i,j,t-1}) \quad (8.5)$$

is the two-period average share for input $X_{i,j}$

The assumption of constant returns to scale implies $\sum_i v_{i,j} = 1$ and allows the use of observed value shares for the calculation of TFP growth in equation (8.3).

Rearranging equation (8.3), yields the standard growth accounting decomposition of output growth into the contribution of each input and the TFP residual

$$\Delta \ln Y_j = \bar{v}_{K,j} \Delta \ln K_j + \bar{v}_{L,j} \Delta \ln L_j + \bar{v}_{E,j} \Delta \ln E_j + \bar{v}_{M,j} \Delta \ln M_j + \bar{v}_{S,j} \Delta \ln S_j + v_{T,j} \quad (8.6)$$

where the contribution of an input is defined as the product of the input's growth rate and its two period average value share.

We can further decompose the contribution of each input into a quantity and quality component. As discussed earlier, the quality component represents substitution between H components, while the quantity component represents the increases in each detailed input. Our extended decomposition of the sources of output growth is given as follows:

$$\Delta \ln Y_j = \bar{v}_{K,j} (\Delta \ln Z_j + \Delta \ln Q_{K,j}) + \bar{v}_{L,j} (\Delta \ln H_j + \Delta \ln Q_{L,j}) + \bar{v}_{E,j} \Delta \ln E_j + \bar{v}_{M,j} \Delta \ln M_j + \bar{v}_{S,j} \Delta \ln S_j + v_{T,j} \quad (8.7)$$

Where Z is the stock of industry capital, Q_k is the capital quality, H is number of persons employed and Q_l is the quality of labour.

The data series at individual industry level constructed for gross value added and gross output, labour input, capital input, intermediate inputs and factor income shares (described in [Chapters 2, 3, 4, 5, 6 and 7](#) respectively) are used to estimate TFPG during the period 1980 to 2019. Labour input is measured by using total person worked along with the labour quality index (see [Chapter 4](#) for a detailed discussion). The series on growth rate in capital services forms the measure of growth in capital input (discussed in [Chapter 5](#)). TFPG estimation is done by using the gross output function framework²¹ (i.e. equation 8.7).

8.2 Methodology for Aggregation across Industries

We also provide the TFPG for aggregate economy and two broad sectors, viz. manufacturing and services.²² However, the method of estimation of TFP growth for individual industries described in section 8.1 cannot be readily applied to a higher level of aggregation. The main problem is that gross value of output cannot be added across industries to generate a measure of output at a higher level of aggregation, say for the economy or of any of the broad sectors. Therefore, it becomes necessary to consider appropriate methods of aggregation across industries consistent with the gross output function specification of technology at the individual industry level.

We use the Törnqvist index²³ for aggregating across the broad sector and total economy. Taking this approach, computations are made for the manufacturing sector, the services sector, and for the

²¹ It should be pointed out that an alternate set of estimates of TFPG obtained by using the value-added function framework is also provided in the India KLEMS dataset for which real gross value added is taken as the measure of output and only the two primary inputs are considered, namely labour input and capital input

²² More sophisticated methods of aggregation, such as the application of production possibility frontier approach or the application of direct aggregation is done in research papers written on the basis of the KLEMS dataset prepared. Such estimates of TFP growth are not a part of the database.

²³ There exists other approaches to aggregate across broad sectors and industry, a detailed discussion on them can be found in KLEMS data manual version 2020 released by RBI on its website

aggregate economy. These aggregates are obtained using Tornqvist aggregate of growth rates of real value added, capital input, and labour input. Below, we briefly describe the Törnqvist index.

Suppose the aggregate sector J consists of n KLEMS industries (j=1:n). We obtain the output (value added) and inputs aggregates for sector J as:

$$\Delta \ln V_J = \sum_j \bar{v}_j \Delta \ln V_j \quad (8.8)$$

$$\Delta \ln K_J = \sum_j \bar{w}_j^K \Delta \ln K_j \quad (8.9)$$

$$\Delta \ln L_J = \sum_j \bar{w}_j^L \Delta \ln L_j \quad (8.10)$$

where \bar{v}_j is nominal value added share, \bar{w}_j^K is the cost share of capital (i.e., the share of rental cost), and \bar{w}_j^L is the labour compensation share – all shares are calculated for KLEMS industry j, in the aggregate sector J, i.e.,

$$v_j = \frac{P_{V,j} V_j}{\sum_j P_{V,j} V_j}; w_j^K = \frac{P_{K,j} S_j}{\sum_j P_{K,j} S_j}; \text{ and } w_j^L = \frac{P_{L,j} H_j}{\sum_j P_{L,j} H_j} \quad (8.11)$$

where $P_{V,j}$ is the price of value added in industry j, V_j is the real value added in industry j, $P_{K,j}$ is the rental price of capital in industry j, S_j is the capital stock in industry j, $P_{L,j}$ is the wage rate (price of labour) in industry j, and H_j is the employment in industry j. The aggregate output (real value added) and input growth rates obtained above are then used with the growth accounting equation to obtain the sectoral TFP growth rates.