Measuring Productivity at the Industry Level
THE INDIA KLEMS DATA BASE

DATA MANUAL 2020
(Version 7)

Deb Kusum Das, Abdul Azeez Erumban, Suresh Aggarwal and Pilu Chandra Das

The India KLEMS research project is housed at the Centre for Development Economics, Delhi School of Economics and is supported by Reserve Bank of India.

AUTHORS

Deb Kusum Das
Department of Economics
Ramjas College
University of Delhi
dkdas@ramjas.du.ac.in

Abdul Azeez Erumban
Groningen Growth and Development Center
Faculty of Economics and Business,
University of Groningen,
a.a.erumban@rug.nl

Suresh Aggarwal
Retired Professor, Department of Business Economics,
University of Delhi, South Campus,
sureshchag@yahoo.com

Pilu Chandra Das
Kidderpore College
University of Calcutta
arpiludas@gmail.com

The Data Manual has been prepared under the guidance of Professor Bishwanath Goldar, Research Advisor, India KLEMS project.
The Authors thank Professor K L Krishna for several suggestions and comments on the manual.
The Authors thank Samiran Dutta for research and editorial support.

Contents

List of Appendixes

List of Tables

List of Box

Changes in India KLEMS Database 2020 over the India KLEMS Database 2019 version

Overall changes

The database is extended to include 2018-19 for all variables, thus providing a series for the period 1980-81 to 2018-19

Variable specific changes

1. **Gross Value Added**
   Estimates of Gross Value Added (GVA) for years since 2011-12, both current and constant (2011-12) prices for all industries are obtained from National Accounts Statistics (NAS) 2021.

2. **Gross value of Output**
   Estimates of Gross Value of Output (GVO) for years since 2011-12, both current and constant (2011-12) prices for all industries are obtained from National Accounts Statistics (NAS) 2021.

3. **Labour Input**

   The latest Periodic Labour Force Survey (PLFS) by NSSO for 2018-19 on employment is used for estimating employment for 2018-19.

4. **Capital Input**

I. Data on investment, capital stock, and capital services for 2018-2019 are added.

II. In the past years, we obtained detailed asset/industry-wise data on capital formation from the CSO. The CSO provided us with two such historical series. The first was in 2004-2005 base, from 1950-51 to 2012-2013, and the second in 2011-2102 base, from 2011-2012 till 2016-2017. Since such data was not available since 2016-2017, the current and previous updates of the India KLEMS rely solely on publicly available data from the NAS for years after 2016.

III. However, given that CSO now provides data that are more detailed to the public, this was largely consistent with our previous data. One outstanding issue was that the asset transport equipment was not separately available. Therefore, we assumed the share of transport equipment in the total machinery and equipment group in 2016-2017 to remain constant in the subsequent years.

IV. The above split is done for current price series. For deflating the assets, asset prices are extrapolated using the trend in the nearest aggregate asset (*e.g.*, transport equipment is deflated using the extrapolated investment price of total machinery).

V. The National Accounts does not provide the split between registered and unregistered manufacturing anymore. Therefore, it was hard to create the control total for these two segments of the manufacturing sector prior to creating the industry series using data from the Annual Survey of Industries and NSSO unorganized manufacturing surveys. In the current version of the database, this issue has been addressed by considering the household sector as a proxy for unregistered and the corporate and public sector as a proxy for the registered sector.

VI. While extending the detailed asset-wise investment series forward using the NAS-published headline series, we 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. This has not been the case in the previous release, where we directly used the 2012-2016 detailed asset series from the CSO. While the current approach is more appropriate and ensures a complete consistency with the NAS published series, it may cause some changes in the capital service growth rates compared to the last release.

VII. The approach to estimating initial capital stock by asset types has been changed. Instead of distributing NAS estimates of net capital stock in the initial year using the asset distribution of GFCF, we now use the industry distribution of net capital stock (see section 5.2d). The new approach, which is more pertinent, leads to changes in aggregate capital stock growth rates in the historical series, as it changes the asset composition of aggregate capital stock in some industries.

VIII. The approach to extrapolate the 2011-2012 base GFCF series backward to 1950 using the 2004-2005 base back series has changed. In the past versions of the database, extrapolation has been made at the detailed NAS industry level, whenever available, which caused some jumps in the data in the bridge year (2011-2012). In the current version, extrapolation has been made at the level of KLEMS industry aggregates (see section 5.2a for detailed discussion.)

IX. This version of the database provides aggregate capital stock at current prices and aggregate capital composition index for the 27 industries.

Chapter 1: Introduction

1.1 Background

This document describes the procedures, methodologies and approaches used in constructing the India KLEMS database version 2020. This database is part of a research project, supported by the Reserve Bank of India (RBI), to analyze productivity performance in the Indian economy at disaggregate industry level. 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.

In its definitive version the India KLEMS research project will include 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. A major advantage of growth accounts is that it is embedded in a clear analytical framework rooted in production functions and the theory of economic growth. It provides a conceptual framework within which the interaction between variables can be analyzed, which is of fundamental importance for policy evaluation (Timmer et.al.2007)[1].

The present document describes the India KLEMS database version 2020. The present version is an extended India KLEMS research project, "Disaggregate Industry Level Productivity Analysis for India- the KLEMS Approach" being undertaken at the Centre for Development Economics, Delhi School of Economics. This one builds on the previous project, which was undertaken at ICRIER, New Delhi[2]. The Data Manual is intended to guide researchers about the variables (and their construction) used to measure both inputs and total factor productivity (TFP) at the industry level using the dataset. In addition, it is also intended to support national officials' statistical agencies in future work on the productivity database within the agencies.

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 2018-19 has been

---

[1] Timmer, M.P, Mahony, M. and van Ark, B. (2007). The EU KLEMS growth and productivity accounts: an overview. Mimeo University of Groningen & University of Birmingham,

[2] Readers can refer to India KLEMS Data Manual Version 2013 for the data set released by RBI on its website in June 2014.

constructed on the basis of data compiled from CSO, 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 a number of 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 2018-19 (2018) 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 so as to generate continuous time series from 1980 to 2018. 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:

- ❑ Agricultural, 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. Further the detailed classification and concordance of study industries with NICs is provided in Appendix table A.

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 Tables (IO). 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. Thus, the use of information from these data sources necessitates interpolation and assumption of constant shares for building series of output and inputs. The construction of growth accounting series like total factor productivity, labour productivity 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 2020

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)

Appendix A: **Concordance table of study industries (minimal) with Different NICs (National Industrial Classifications)**

Sl. No.	KLEMS Code	Description (NIC & KLEMS)	NIC 2008	NIC 2004	NIC 1998	NIC 87	NIC70
1	A to B	Agriculture, Forestry and Fishing	0111+0112+0113+0114+0115+0116+0119+0121+0122+0123+0124+0125+0126+0127+0128+0129+0130+0141+0142+0143+0144+0145+0146+0149+0150+0161+01+62+01631+01633+01639+0164+0170+0210+0220+0230+0240+0311+0312+0321+0322	0111+0112+0113+0121+0122+0130+01401+01402+01403+01404+01406+01407+01408+01409+0150+0200+0501+0502	0111+0112+0113+0121+0122+0130+01401+01402+01403+01404+01406+01407+01409+0150+0200+0500	000+001+002+003+004+005+006+008+009+012+013+018+007+010+011+014+015+016+017+019+020+021+022+023+024+025+026+027+029+030+031+032+033+034+035+036+037+039+040+050+051+052+053+054+059+060+061+062+063+064+069	000+001+002+003+004+005+006+008+012+013+018+007+010+011+014+015+016+017+019+020+021+022+023+024+025+026+029+030+031+032+033+034+035+036+039+040+050+051+052+053+054+059+060+061+062+063+069
2	C	Mining and Quarrying	0510+0520+0610+0620+0710+0721+0729+0810+0891+0892+0893+0899+0910+0990	1010+1020+1030+1110+1120+1200+1310+1320+1410+1421+1422+1429	1010+1020+1030+1110+1120+1200+1310+1320+1410+1421+1422+1429	100+101+102+110+111+190+140+120+130+131+132+133+134+135+136+137+138+139+150+151+152+153+154+155+156+159	100+101+110+111+120+121+122+123+124+125+126+127+128+129+190+191+192+193+194+195+199
3	15 to 16	Food and Beverages and Tobacco	1010+1020+1030+1040+1050+1061+1062+1071+1072+1073+1074+1075+1079+1080+1101+1102+1103+1104+1200	1511+1512+1513+1514+1520+1531+1532+1533+1541+1542+1543+1544+1549+1551+1552+1553+1554+1600	1511+1512+1513+1514+1520+1531+1532+1533+1541+1542+1543+1544+1549+1551+1552+1553+1554+1600	200+203+202+210+211+212+201+204+218+217+205+206+207+209+213+214+215+219+220+223+221+222+216+224+225+226+227+228+229	200+203+202+210+211+212+201+204+217+205+206+207+209+213+214+215+219+220+223+221+222+216+224+225+226+227+228+229+315

4	17 to 19	Textiles, Textile Products and Leather and Footwear	1311+1312+1313+1391+1392+1393+1394+1399+14101+14102+14103+14104+14109+1420+1430+1511+1512+1520+01632	1711+1713+1712+1714+1721+1722+1725+1723+1724+1729+1730+18101+18102+18103+18104+18109+1820+1911+1912+1920+01405	1711+1712+1721+1722+1723+1729+1730+18101+18102+18103+18104+18109+1820+1911+1912+1920+01405	230+231+232+233+234+235+240+241+242+244+245+247+250+251+252+253+254+255+256+236+243+246+248+257+258+259+267+268+263+264+261+262+269+260+265+266+292+294+295+296+290+293+299+291+311	230+231+232+233+234+235+236+240+241+242+243+245+246+247+248+244+250+251+252+253+259+260+261+262+263+264+265.2+265.3+266+267+268.1+268.2+269+290+291+292+293+294+295+296+299+301+239+249
5	20	Wood and Products of Wood and Cork	1610+1621+1622+1623+1629	2010+2021+2022+2023+2029	2010+2021+2022+2023+2029	270+271+272+273+274+275+279	271+270+273+272+274+275+279
6	21 to 22	Pulp, Paper and Paper Products and Printing and Publishing	1701+1702+11709+1811+1812+1820+5811+5812+5813+5819	2101+2102+2109+2211+2212+2213+2219+2221+2222+2230	2101+2102+2109+2211+2212+2213+2219+2221+2222+2230	280+281+282+283+285+284+286+289+287+288	280+281+282+283+285+284+286+289+287+288
7	23	Coke, Refined Petroleum and Nuclear Fuel	1910+1920	2310+2320+2330	2310+2320+2330	318+319+314+315+316+317	304+305+306+307
8	24	Chemicals and Chemical Products	2011+2012+2013+2021+2022+2023+2029+2030+2100+2680	2411+2412+2413+2421+2422+2423+2424+2429+2430	2411+2412+2413+2421+2422+2423+2424+2429+2430	300+301+302+303+304+305+208+307+308+309+306	208+310+311+312+313+314+316+317+318+319
9	25	Rubber and Plastics	2211+2219+2220	2511+2519+2520	2511+2519+2520	310+312+313	300+302+303
10	26	Other Non-Metallic Mineral	2310+2391+2391+2394+2395+2396+2399	2610+2691+2692+2693+2694+2695+2696+2699	2610+2691+2692+2693+2694+2695+2696+2699	321+322+323+320+324+327+326+325+329	320+321.1+321.2+321.3+321.4+321.6+321.7+321.9+322+323+324+325+326+327+328+329

11	27 to 28	Basic Metals and Fabricated Metal Products	2410+2420+2431+2432+2511+2512+2513+2591+2592+2593+2599	2711+2712+2713+2714+2715+2716+2717+2718+2719+2720+2731+2732+2811+2812+2813+2891+2892+2893+2899	2710+2720+2731+2732+2811+2812+2813+2891+2892+2893+2899	330+331+332+333+334+335+336+337+338+339+340+341+343+344+345+346+349+352	330+331+332+333+334+335+336+339+340+341+343+344+345+349+352
12	29	Machinery, nec	2520+2750+2811+2812+2813+2814+2815+2816+2817+2818+2819+2821+2822+2823+2824+2825+2826+2829+3040+3311+3312	2911+2912+2913+2914+2915+2919+2921+2922+2923+2924+2925+2927+2929+2930	2911+2912+2913+2914+2915+2919+2921+2922+2923+2924+2925+2927+2929+2930	355+364+388+350+351+353+354+356+357+359+390+391+392+393+397+399	350+351+353+354+356+357+359+355+363.1+363.2+363.5+363.6+363.7+363.8+363.9
13	30 to 33	Electrical and Optical Equipment	2610+2620+2630+2651+2652+2660+2670+2710+2720+2731+2732+2733+2740+2790+3250+3314+3319+3320+9512+9521	3000+3110+3120+3130+3140+3150+3190+3210+3220+3250+3311+3312+3313+3320+3330	3000+3110+3120+3130+3140+3150+3190+3210+3220+3250+3311+3312+3313+3320+3330	358+367+360+395+361+362+363+369+368+365+396+366+380+381+382	321.5+358+360+361+362+363.3+363.4+364+365+366+367+369+380+381+382
14	34 to 35	Transport Equipment	2910+2920+2930+3011+3012+3020+3030+3091+3092+3099+3315	3410+3420+3430+3511+3512+3520+3530+3591+3592+3599	3410+3420+3430+3511+3512+3520+3530+3591+3592+3599	373+374+370+371+372+377+375+376+378+379	373+374+370+371+372+377+375+376+378+379
15	36 to 37	Manufacturing nec, Recycling	3100+3211+3212+3220+3230+3240+3290+3830	3610+3691+3692+3693+3694+3699+3710+3720	3610+3691+3692+3693+3694+3699+3710+3720	276+277+342+383+384+386+385+387+389	265.1+276+277+342+383+384+385+386+387+389
16	E	Electricity Gas and Water Supply	3510+3520+3530+3600	4010+4020+4030+4100	4010+4020+4030+4100	400+401+430+431+432+439+410+420	400+401+410+420
17	F	Construction	4100+4210+4220+4290+4311+4312+4321+4322+4329+4330+4390	4510+4520+4530+4540+4550	4510+4520+4530+4540+4550	191+199+500+501+502+503+504+505+506+509+510+511+514+519+512+513+515	500+501+502+503+504+505+509+510+511+512+513+514+519

18	50 to 52	Trade	4510+4520+4530+4540+4610+4620+4630+4641+4649+4651+4652+4653+4659+4730+9200+4711+4719+4721+4722+4723+4741+4742+4751+4752+4753+4759+4761+4762+4763+4764+4771+4772+4773+4774+9522+9523+9529	5010+5020+5030+5040+5050+5110+5121+5122+5131+5139+5141+5142+5143+5149+5151+5152+5159+5190+5211+5219+5220+5231+5232+5233+5234+5239+5240+5251+5252+5259+5260	5010+5020+5030+5040+5050+5110+5121+5122+5131+5139+5141+5142+5143+5149+5150+5190+5211+5219+5220+5231+5232+5233+5234+5239+5240+5251+5252+5259+5260	623+686+398+974+682+640+641+642+649+600+603+604+605+601+606+607+608+609+612+611+615+617+618+622+630+631+634+636+637+639+613+614+619+610+632+633+602+616+638+620+621+635+684+685+840+688+841+650+651+652+653+654+655+656+659+687+660+661+671+672+673+674+675+679+683+670+676+680+681+689+970+971+972+973+975+979	632+683+973+600+601+602+603+604+605+606+607+608+610+611+612+613+620+621+622+623+630+631+640+641+642+643+644+645+646+647+649+682+821+650+651+652+653+654+655+656+659+660+661+670+671+672+673+674+675+676+679+681+684+680+689+971+972+974+975+979
19	H	Hotels and Restaurants	5510+5520+5590+5610+5621+5629+5630	5510+5520	5510+5520	691+690	691+690
20	60 to 63	Transport and Storage	4911+4912+4921+4922+4923+4930+5011+5012+5021+5022+5100+5120+5210+5229+7911+7912+7990+5211+5222+5223+5224	6010+6021+6022+6023+6030+6110+6120+6210+6220+6301+6302+6303+6304+6309	6010+6021+6022+6023+6030+6110+6120+6210+6220+6301+6302+6303+6304+6309	700+701+702+709+703+704+705+706+707+710+711+720+730+731+732+740+741+749+708+712+721+737+738+739	700+701+702+703+704+705+706+707+708+710+711+712+720+721+730+740+741+749
21	64	Post and Telecommunications	5310+5320+6110+6120+6130+6190	6411+6412+6420	6411+6412+6420	750+751+752+759	750+751+759
22	J	Financial Intermediation	6411+6419+6420+6430+6491+6492+6499+6511+6512+6520+6530+6611+6612+6619+6621+6622+6629+6630+7740	6511+6519+6591+6592+6599+6601+6602+6603+6711+6712+6719+6720	6511+6519+6591+6592+6599+6601+6602+6603+6711+6712+6719+6720	800+801+802+803+804+811+810+819+812	800+801+809+811+810+819

23	71-74	Business Services	6201+6202+5820+6209+6311+6312+6339+6910+6920+7010+7020+7110+7210+7310+7410+7120+7220+7320+7420+7490+7710+7721+7722+7729+7730+7810+7820+7830+8010+8020+8030+8110+8121+8129+8211+8219+8220+8230+8291+8292+8299+8550+9511	7111+7112+7113+7121+7122+7123+7129+7130+7210+7221+7229+7230+7240+7250+7290+7310+7320+7411+7413+7414+7421+7422+7430+7491+7492+7493+7494+7495+7499	7111+7112+7113+7121+7122+7123+7129+7130+7210+7220+7230+7240+7250+7290+7310+7320+7411+7413+7414+7421+7422+7430+7491+7492+7493+7494+7495+7499	733+734+850+736+735+851+853+852+854+892+394+922+830+891+893+894+895+896+898+890+899+963	822+823+824+825+826+827+829+830+922+963
24	L	Public Admin and Defence; Compulsory Social Security	8411+8412+8413+84211+8422+8423+8430	7511+7512+7513+7514+7521+7522+7523+7530	7511+7512+7513+7514+7521+7522+7523+7530	900+901+902+903	900+901+902+903
25	M	Education	8510+8521+8522+8530+8541+8542+8549	8010+8021+8022+8030+8090	8010+8021+8022+8030+8090	921+920	921+920
26	N	Health and Social Work	8610+8620+8690+8710+8720+8730+8790+8810+8890+7500	8511+8512+8519+8520+8531+5520+8532	8511+8512+8519+8520+8531+8532	930+931+941	930+931+941
27	O	Other Services	6810+6820+9311+9312+9319+9321+9329+9411+9412+9420+9491+9492+9499+9601+9602+9603+9609+9700+9810+9820+9900+3811+3812+3821+3822+3900+3700+5911+5912+5913+5914+9000+6391+6010+8541+14105+6020+9101+9102+9103+8130+5920	7010+7020+9000+9111+9112+9120+9191+9192+9232+9199+9211+9212+9213+9214+9219+9220+9231+9233+9241+9249+9301+9302+9303+9309+9500+9600+18105	7010+7020+9000+9111+9112+9191+9199+9211+9212+9213+9214+9219+9120+9192+9220+9231+9232+9233+9241+9249+9301+9302+9303+9309+9500+18105	820+910+942+940+943+949+950+951+954+952+953+955+958+897+957+956+959+961+962+969+990+960 + 980+964	820+828+910+940+942+949+950+951+952+953+954+955+956+959+961+962+969+990+960

***Source**: Prepared by authors using Central Statistical Office (CSO) reports, India*

****

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

For an individual firm or industry, productivity measure can be based on a value-added concept where value added is considered as an industry's output and only primary inputs such as labour and capital are considered as industry inputs. Value added based productivity measures reflect an industry's capacity to contribute to economy-wide income and final demand. In this sense, they are valid complements to gross output-based measures. 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 2018-19 (2018).

2.1 Methodology

GVA of a sector is defined as the value of output less the value of its intermediary inputs. This value added created by a sector is shared among the primary factors of production, labour and capital. The National Accounts Statistics (NAS) brought out by the CSO (Central Statistics Office, Government of India) is the basic source of data for the construction of series on gross value added for INDIA KLEMS-industries. NAS provides estimates of GVA (*i.e.*, gross value added) at a disaggregated industry level at both current and constant prices for the period since 1950-51. 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 2021. 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.

In the previous versions of the India KLEMS data, value-added was measured at factor cost (as was the case with the NAS). However, this was not fully consistent with the KLEMS standards used internationally, as the international standards of KLEMS methodology were to follow value added at basic prices.[3] Our choice to use the factor price concept was driven by

[3] The difference between value added at factor cost and value added at basic price is the treatment of taxes and subsidies on production. While the former excludes all taxes on production and includes all subsidies (on intermediate inputs and factor inputs), the latter includes only taxes and subsidies on intermediate inputs. In practice, gross value added at basic prices is measured as output valued at basic prices less intermediate consumption valued at purchasers' prices. The basic price is the amount receivable by the producer from the purchaser for a unit of a good or service produced as output minus any tax payable, and plus any subsidy receivable, on that unit as a consequence of its production or sale. It excludes any transport charges invoiced separately by the producer. Although the outputs and inputs are valued using different sets of prices, for brevity the value added is described by the prices used to value the outputs. From the point of view of the producer, purchasers' prices for inputs and basic prices for outputs represent the prices actually paid and received. Their use leads to a measure of gross value added which is particularly relevant for the producer.
Gross value added at factor cost can easily be derived from the measures of gross value added at basic prices presented above by subtracting the value of any taxes, less subsidies, on production payable out of gross value added as defined. For example, the only taxes on production remaining to be paid out of gross value added at basic prices consist of "other taxes on production". These consist mostly of current taxes (or subsidies) on the labour or capital employed in the enterprise, such as payroll taxes or current taxes on vehicles or buildings. Gross value added at factor cost can, therefore, be derived from gross value added at basic prices by subtracting "other taxes, less subsidies, on production.

the fact that the NAS provided gross value added at current and constant prices only at factor cost. Since 2011-2012, following the SNA, the NAS started providing value added at basic prices, which we could directly use and keep the India KLEMS approach consistent with the general KLEMS standards. Therefore, the new version of the India KLEMS database bypasses the inconsistency with the KLEMS standards we had while maintaining consistency with the NAS, as the NAS has now entirely shifted to the basic price concept.

NAS provides separate estimates of GVA for registered and unregistered manufacturing. However, onwards 2011-12 NAS has disaggregated the manufacturing sector in to corporate sector and household sector. For splitting the aggregate estimates of GVA for registered manufacturing industries and the corporate sector, we have used data from the Annual Survey of Industries (ASI) based on the National Industrial Classification 2004 and 2008 (NIC-2004 & NIC-2008). Whereas, for the unregistered manufacturing sector, we have used results from six rounds of NSSO surveys- [40th (1984-85), 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)] to obtain the value-added estimates. In India, GDP for unregistered manufacturing is constructed using the labour input method. The estimates of GVA for the unregistered manufacturing sector are obtained as a product of the workforce and the corresponding GVA per worker. The information about employment in the unorganized sector is only available in the benchmark years for which NSSO survey data are available. Therefore, there is no consistent source of employment data for the years between these quinquennial surveys. Even the information on value added per worker is equally limited since the value-added data are also updated on an approximate 5-year interval (for details, see CSO, 2007). Therefore, estimates of value added for the unregistered manufacturing sectors for the years between the benchmarks have been obtained by interpolation and for years outside the benchmark years by linear extrapolation. For splitting the aggregate estimates of GVA for the household sector for recent years (*i.e.*, 2011 onwards), we have used GVA data from 67th round (2010-11) and 73rd round (2015-16).

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 available from NAS[4]. The sectors for which data are provided in NAS are Agriculture, Forestry & logging, Fishing, Mining and Quarrying, Manufacturing (registered and un-registered), Electricity, Construction, Trade, Hotels & Restaurants, Railways, Transport by other means, Storage, Communication, Banking & insurance, Real estate,

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

Ownership of Dwelling & Business Services, Pubic 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. GVA data are directly available from NAS for 6 out of 13 manufacturing sectors the period 1980-81 to 2011-12, while for 8 out of 13 manufacturing sectors GVA data directly available since 2011-12. The list of these industries is provided in the table below.

Table 2. 1: List of Manufacturing Industries for which GVA data is directly available from NAS

1980-81 to 2011-12		Since 2011-12	
Industry No.	**Industry Description**	**Industry No.**	**Industry Description**
3	Food and Beverages and Tobacco	**3**	Food Products, Beverages and Tobacco
4	Textiles, Textile Products, Leather and Footwear	**4**	Textiles, Textile Products, Leather and Footwear
6	Pulp, Paper, Paper Products, Printing and Publishing	**5**	Wood and Products of wood
8	Chemicals and Chemical Products	**6**	Pulp, Paper, Paper products, Printing and Publishing
10	Other Non-Metallic Mineral Products	**7**	Coke, Refined Petroleum Products and Nuclear fuel
14	Transport Equipment	**8**	Chemicals and Chemical Products
		9	Rubber and Plastic Products
		10	Other Non-Metallic Mineral Products

Source: *National Account Statistics (NAS) reports.*

For the remaining industries GVA data is constructed by splitting the NAS data using ASI or NSSO distributions. ASI data (annual) has been used for registered 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 Unregistered Manufacturing segments. A list of study industries is presented in Table 2.2 and 2.3 showcasing the methodology used to split GVA of certain NAS sectors to match concordance with our classification.

Table 2. 2: 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 *etc.* (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*

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

Following are the details of steps taken in splitting NAS sectors into India KLEMS industries for which direct Gross Value Added series are not available between 1980 and 2011.

Wood and wood products and manufacturing of furniture

In India KLEMS the sectors Wood products (20) and Furniture and fixtures (361) are considered separately, with the latter being added to the sector Manufacturing nec and Recycling (36+37). These sectors are available separately since 2011-12 from National Accounts (NAS 2021). However, the Back Series of National Accounts (Base 2011-12) does not provide them separately; rather it provides the data for the aggregate sector Wood and

Wood products, furniture, fixtures *etc.* (20+361) – both for registered and unregistered sectors. For the registered segment of the industry, we split the value added in this aggregate sector to two individual sectors using the share of the respective sectors from the Annual Survey of Industries (ASI). For the unorganized segment, we use the shares of the respective sectors obtained from NSSO unorganized manufacturing surveys of the benchmark years (1984-85, 1989-90, 1994-95, 2000-01, 2005-06 and 2010-11), to split the aggregate data. The shares for the interim years between two benchmark years have been linearly interpolated till 2011-12, and for the period 1980-81 to 1984-85, the share of 1984-85 has been used.

Coke, Refined Petroleum Products and Nuclear Fuel and Rubber and Plastic Products

We split 'Rubber, Petroleum Products' (which are clubbed under one group in NAS) to arrive at two industry groups *i.e.*, Coke, Refined Petroleum Products and Nuclear Fuel (23) and Rubber and Plastic Products (25). For the organized segment, we use the ASI (annual) data to get the individual sector shares and split the NAS data using these individual shares. Likewise, we use the relevant data from the four NSS surveys mentioned earlier to get the individual sector shares for the unorganized segment of this sector. The ratio of GVA for the interim years between two benchmark years have been linearly interpolated till 2011-12, and from 1980-81 to 1984-85, the ratio of 1984-85 has been used.

Basic Metals and Fabricated Metal Products

In our industry classification, Basic Metals and Fabricated Metal products (27+28) and Machinery (29) are separate groups whereas 'Fabricated Metal Products' (28); 'Machinery and Equipment nec' (29) and 'Office, Accounting and Computing Machinery' (30) are clubbed together as Metal Products and Machinery (28, 29 and 30) in NAS. To arrive at individual industry result, we use ASI shares for organized sectors and NSSO surveys for the unorganized sector. We add to the fraction of fabricated metal products (28) from Metal Products and Machinery to basic metals (271+272+2731 +2732) already available.

Electrical and Optical Equipment

In our study classification, 'Electrical and Optical equipment' includes all sectors from 30 to 33 under NIC 1998. However, 'Electrical Machinery' in NAS includes industries 'Electrical Machinery and Apparatus n.e.c.' (31) + 'Radio, Television and Communication Equipment and Apparatus' (32) and excludes 'Office, Accounting and Computing Machinery' (30) and 'Medical, Precision and Optical Instruments' (33). However, 30 is part of 'Metal Products and Machinery' and 33 is part of 'Other Manufacturing' in NAS. We take out 28 from Metal Products and Machinery in NAS, with 29 and 30 being left, which we split using ASI. NSSO surveys have been useful here as well to compute the unorganized segment share. Likewise, we also take out the share of 33 from 'Other Manufacturing' in NAS separately for both organized and unorganized segments to arrive at gross value added for electrical and optical equipment.

NAS provides estimates of GVA for more disaggregate industries at both current and constant prices for the period since 2011-12. For 8 out of 13 manufacturing sectors GVA data are directly available from NAS. For the remaining 5 industries GVA data is constructed by splitting the NAS data using ASI or NSSO distributions.

Table 2. 3: List of Manufacturing Industries for which Gross Value-Added data is obtained by adjusting data for NAS Industries (Onwards 2011)

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*

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

Following are the details of steps taken in splitting NAS sectors into India KLEMS industries for which direct Gross Value Added series are not available since 2011-12.

Basic Metals and Fabricated Metal Products

In our industry classification, Manufacture of structural metal products, tanks, reservoirs and steam generators (251) and Manufacture of other fabricated metal products; metalworking service activities (259) are combined as ‘Fabricated Metal Products’ (251+259) and Manufacture of weapons and ammunition (252) included in ‘Machinery nec’ whereas 251, 252 and 259 are clubbed together as Manufacture of fabricated metal products, except machinery and equipment (25) in NAS. To arrive at individual industry result, we use ASI shares for organized sectors and NSSO surveys for the unorganized sector. We add to the fraction of fabricated metal products (251 & 259) to basic metals (241+242+2431+2432) already available.

Machinery, nec.

In our study classification, ‘Machinery, nec.’ Includes 28, 252, 275, 304, 3311 and 3312 under NIC 2008. However, in NAS, Manufacture of fabricated metal products, except machinery and equipment (25) includes Manufacture of weapons and ammunition (252), Manufacture of electrical equipment (27) includes Manufacture of domestic appliances (275), Manufacture of transport equipment (29+30) includes Manufacture of military fighting vehicles (304), Repair and installation of machinery and equipment (33) includes Repair of machinery (3311 & 3312). We take out the required portion, which we split using ASI & NSSO ratio and added to already available Manufacture of machinery and equipment n.e.c. (28).

Electrical and Optical Equipment

In our study classification, ‘Electrical and Optical equipment’ includes 26, 27-275, 325, 3313, 3314, 3319, 3320 under NIC 2008. Using ASI and NSSO portion, we excluded the unrequired portion Manufacture of domestic appliances (275) from Manufacture of electrical equipment (27), and take out the required portion Manufacture of medical and dental instruments and supplies (325) from Other manufacturing (32) and Repair of electronic and optical equipment, electrical equipment and other equipment (3313, 3314, 3319 & 3320) from Repair and installation of machinery and equipment (33).

Transport Equipment

We excluded ‘Manufacture of military fighting vehicles (304) from Manufacture of transport equipment (29+30) and included Repair of transport equipment (3315) to arrive at industry group - Transport Equipment.

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

Step 3: According to India KLEMS, output is adjusted for Financial Intermediation Services Indirectly Measured (FISIM). The value of such services forms a part of the income originating in the banking and insurance sector and, as such, is deducted from the GVA. The NAS provides output net of FISIM for some industry groups at a more aggregate level. For instance, in the estimates of GVA obtained for the registered manufacturing sector, adjustment for FISIM in NAS is made only at the aggregate level in the absence of adequate details at a disaggregate level. However, we have allocated FISIM to all the sectors of manufacturing by redistributing total FISIM across sectors proportional to their sectoral GDP shares. Similar redistribution of FISIM has been done in case of Real Estate, Ownership of Dwellings and Professional Services; and Other Services sector.

****

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 2018-19(2018). We discuss both the raw data sources and the adjustments that have been made to generate the time series on output and value added consistent with the official National Accounts. The methodology for measuring industry output, and value added was developed by Jorgenson, Gallop and Fraumeni (1987) and extended by Jorgenson (1990 a). Following a similar approach as explained in Jorgenson *et al.* (2005, Chapter 4) and Timmer *et al.* (2010, Chapter 3), the time series on gross output and intermediate inputs for the Indian economy have been constructed.

3.1 Methodology

The gross output of an industry is defined as the value of industry production using primary factors like labour, capital and intermediate inputs purchased from other industries. The gross output production function assumes separability in inputs and technology. An important advantage of the gross output approach is that it provides a complete measure of production and treats all inputs - labour, capital and intermediate inputs symmetrically. In contrast, the value-added measure of output does not explicitly account for the flow of intermediate inputs which may be the primary component of an industry's output. We use the more restrictive value-added concept primarily because it is useful for aggregation purposes. It is to be noted that aggregate output (aggregated over industry value added), is a value-added concept and the detailed methodology of aggregation of output across industries is explained in chapter 8.

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 source 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 2020, 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 2021.

(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.

Annual Survey of Industries and NSSO Quinquennial Survey Reports:

As mentioned above, gross output data are available at a more disaggregated level in Annual Survey of Industries (ASI) and NSSO quinquennial surveys for Registered and Unregistered Manufacturing industries, respectively. These secondary data sources are used in this study for two purposes: (a) in certain cases NAS provides combined estimates of GVO and GVA for two manufacturing industries. In such cases separate estimates for individual study industries are obtained with the help of ASI or NSSO unorganized manufacturing sector data. (b) For the period prior to 2004, NAS does not provide estimates of GVO for unorganized manufacturing industries. To make our estimate of GVO for this period, the NSSO data are used. The Major NSSO Rounds for Unregistered Manufacturing used are 40th Round (1984-85), 45th Round (1989-90) and 51st Round (1994-95), 56th Round (2000-01), 62nd Round (2005-06), 67th Round (2010-11) and 73rd Round (2015-16).

Input Output Transaction Tables:

As motioned 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 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 from 1980 to 2017 at current and constant prices involves the following steps:

Step 1: Measuring Gross Output of Agricultural Sector, of 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 1980-2018 is also directly taken from NAS.

Step 2: Measuring Gross Output of Manufacturing Industries

For manufacturing industries time series on gross output is obtained by adding the magnitudes for registered and unregistered segments of manufacturing. As mentioned earlier, NAS estimates of gross output for manufacturing industries are at a more aggregate level. In such cases the aggregate output of NAS at current prices has been split using additional information from ASI and NSSO unorganized sector reports. Since 2011-12, gross output data for 8 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 5 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 2018.

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

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*

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

The detailed method of splitting the output of NAS sectors to derive output of individual industries is given as follows:

Basic Metals and Fabricated Metal Products

In our industry classification, Manufacture of structural metal products, tanks, reservoirs and steam generators (251) and Manufacture of other fabricated metal products; metalworking service activities (259) are combined as ‘Fabricated Metal Products’ (251+259) and Manufacture of weapons and ammunition (252) included in ‘Machinery nec’ whereas 251, 252 and 259 are clubbed together as Manufacture of fabricated metal products, except machinery and equipment (25) in NAS. To arrive at individual industry result, we use ASI shares for organized sectors and NSSO surveys for the unorganized sector. We add to the fraction of fabricated metal products (251 & 259) to basic metals (241+242+2431+2432) already available.

Machinery, nec.

In our study classification, 'Machinery, nec.' Includes 28, 252, 275, 304, 3311 and 3312 under NIC 2008. However, in NAS, Manufacture of fabricated metal products, except machinery and equipment (25) includes Manufacture of weapons and ammunition (252), Manufacture of electrical equipment (27) includes Manufacture of domestic appliances (275), Manufacture of transport equipment (29+30) includes Manufacture of military fighting vehicles (304), Repair and installation of machinery and equipment (33) includes Repair of machinery (3311 & 3312). We take out the required portion, which we split using ASI & NSSO ratio and added to already available Manufacture of machinery and equipment n.e.c (28).

Electrical and Optical Equipment

In our study classification, 'Electrical and Optical equipment' includes 26, 27-275, 325, 3313, 3314, 3319, 3320 under NIC 2008. Using ASI and NSSO portion, we excluded the unrequired portion Manufacture of domestic appliances (275) from Manufacture of electrical equipment (27), and take out the required portion Manufacture of medical and dental instruments and supplies (325) from Other manufacturing (32) and Repair of electronic and optical equipment, electrical equipment and other equipment (3313, 3314, 3319 & 3320) from Repair and installation of machinery and equipment (33).

Transport Equipment

We excluded 'Manufacture of military fighting vehicles (304) from Manufacture of transport equipment (29+30) and included Repair of transport equipment (3315) to arrive at industry group - Transport Equipment.

The nominal estimates of output for manufacturing sectors are then deflated with suitable WPI deflators to arrive at the constant price series.

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

Onwards 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 has been constructed using information from Input – Output Transaction Tables of the Indian economy published by CSO.

- GVO to GVA ratios for Services sectors are obtained from IOTT benchmark years of 1978-79, 1983-84, 1989-90, 1993-94, 1998-99, 2003-04, 2007-08.
- These ratios are linearly interpolated for intervening years and applied to GVA series of NAS to derive the output estimates consistent with NAS both at current and constant prices.
- It is to be noted that for government owned sector Public Administration and Defense no intermediate inputs are given in IOTT tables. Consequently, value added to output ratio from The System of National Accounts (SNA) tables have been applied to real and nominal GVA figures of NAS to estimate the output for this sector.

3.3 Outstanding Issues

Firstly, the present study provides estimates for manufacturing and its sub branches without segregating manufacturing (and its sub branches) into organized and unorganized segments. However, given the employment potential and sizable presence of the unorganized segment in many of the manufacturing industries, it would be worthwhile in Indian context to examine separately the productivity performances of both the organized and unorganized components. Some work to construct the output and input series separately for organized and unorganized components of Indian manufacturing has been done. A paper based on this analysis has been prepared. The data series for organized and unorganized manufacturing is not included in this release of India KLEMS database.

Finally, National Accounts do not provide any estimates of gross output of services sector and hence we rely on Input output transaction tables which are available at an interval of about 5 years. This necessitates interpolation and assumption of constant shares for measuring output of services sectors. This issue is analogous to those explained in Timmer *et al.* (2010, Chapter 3) for the EU economy. Griliches (1994) paid particular attention to service sector output as a key source of uncertainty.

****

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 2018. Labour input is measured by combining data on labour persons and data on labour quality. In the KLEMS framework it is desirable to estimate changes in labour quality by industries on the basis of age, gender and education. The measurement of labour quality is essentially an attempt to distinguish one labour type from the other taking into account the embodied human capital in each person. The source of human capital could be through investment in education, experience, training, *etc*. The contribution to output by each person also comes from this embodied capital and 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. Nevertheless, many limitations of India's employment statistics, 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. The problems of employment statistics in India have been widely discussed in the literature (Sivasubramonian; 2004, & Himanshu; 2011, Ghose; 2016). This study aims to build a time series of employment series for 27 industrial sectors. However, there exists no time-series data on Indian economy, except for the organized segment. Therefore, it was essential to make certain assumptions regarding the annual changes in the employment series using available information.

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 and 2018-19 by National Statistical Organisation (NSO) are the main sources for estimating the total workforce in the country by industry groups, as per the National Industrial Classification (NIC). The estimates obtained from EUS are adjusted for population. The major round 32nd (1977-78) has been used for extrapolating the labour series to 1980-81[5].

The source of data for the population is Census of India (different years), which only gives decadal census, so interpolated mid-year population is used for intervening years.

Methodology

a) *Measuring persons employed at the Industry Level*

Efforts are made in this section to describe the methodology used to estimate persons employed by 27 industry and for broad sectors and adjust that measure for changes in labour skill by calculating the labour quality index, thus obtaining the quality corrected labour input.

Since NSSO uses National Industrial Classification 1970 (NIC) for classification of persons employed by industry in 38th and 43rd rounds, NIC 1987 for 50th round, NIC 1998 for 55th and 61st rounds, and NIC 2008 for 68th round and PLFS, therefore as a starting point *concordance between the 27 sectors, and NIC-1970, 1987, 1998 and 2008 was done.*

The underlying methodological issue is how to estimate number of persons employed. In India, the total workforce in the country and its distribution over economic activities may be obtained from the decennial Population Census and the Employment and Unemployment Surveys (EUS) of the NSSO[6]. Out of the two, the latter are more dependable and have been used to assess the changes in employment and unemployment for employment planning and policy analysis. The preference for the use of EUS is generally based on the notion that prior to 2001, the three Censuses have clearly under reported the participation of women in economic activities; whereas the EUS has provided reasonably reliable estimates of the level and pattern of employment (Visaria; 1996). While Population Census underestimates work force participation rates (WPRs), the EUS estimates of total population are significantly lower than the Population Census based estimates – by over 20 percent in Urban India[7]. However, for the Census 2001, the WFPRs are closer to the rates from the 1999-2000 NSSO round. Due to these advantages of EUS, the present study has used the major rounds of EUS and the PLFS for the years till 2018-19. Thus, the surveys used are 38th, 43rd, 50th 55th, 61st, 68th rounds and PLFS (2017-18 and 2018-19)[8].

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.[9] UPSS is the most liberal and

---

[5] Refer to Appendix B for details about all the employment and unemployment surveys conducted by NSSO in India.

[6] This procedure has been advocated by NSSO itself and has been consistently adopted by almost all researchers.

[7] A very lucid description of the same is given by Sivasubramonian (2004).

[8] The 66th round has not been used because of its inconsistency with the 61st and 68th round. Most economists consider it as an outlier.

[9] Refer to Appendix C for detailed definition of UPSS, CWS and CDS.

widely used of these concepts and despite that the UPSS has some limitations[10], this seems to be the best measure to use given the data.

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

- Work Participation Rates (WPRs) by UPSS of the four components - Rural Male, Rural Female, Urban Male and Urban Female from EUS are applied to the corresponding period's census population to find out the number of persons employed in the four segments and the total employment (Appendix Table D)[11]
- 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
- The estimates of employment for the intervening years between the major EUS rounds and PLFS 2017-18 have been obtained by simple interpolation of the four segments- rural male, rural females, urban male and the urban females, ignoring the outliers where the two closest rounds have then been used for interpolation. While, the estimates from the major rounds are centred around the mid-year of the survey rounds, i.e July for 38th round and 1st Jan for the other rounds; the interpolation has been centred around 1st October so as to coincide with the mid-year of the national income estimates which refer to the financial year. So, the base is shifted to October of each year while interpolating the values.
- Total persons in a year were then obtained for each industry as the sum of the Lij over gender and sectors, *i.e.*, $\Sigma_i\Sigma_j L_{ij}$

For extrapolation backward to 1980-81 to 1982-83, the interpolation of the broad industrial classification of 32nd round (1977-78) and 38thround (1983-84) is used. Thus, the estimates from 32nd round are mainly used as control numbers.

Between the survey periods interpolation of estimates was done to find employment for the intervening period. However, in few survey periods where the disaggregate employment of an industry was an outlier, it was ignored for interpolation and the interpolation in such cases was done between the other two adjacent rounds. 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

---

[10] Problems in using UPSS are: The UPSS seeks to place as many persons as possible under the category of employed by assigning priority to work; no single long-term activity status for many as they move between statuses over a long period of one year, and Usual status requires a recall over a whole year of what the person did, which is not easy for those who take whatever work opportunities they can find over the year or have prolonged spells out of the labour force.

[11] Appendix Table D provides the details for the four segments about the Census Population, WFPR by UPSS, and persons employed in each of the EUS rounds used in the study.

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 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 in to 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 this method the aggregate labour input L_j of sector 'j' is defined as a Törnqvist volume index of persons worked by individual labour types 'l' as follows:[12]

$$\Delta \ln L_{j,t} = \sum \bar{v}_{l,j,t} \Delta \ln H_{l,j,t} \quad (4.1)$$

with weights given by

$$v_{l,j}^{L} = \frac{P_{l,j}^{L} \cdot H_{l,j}}{\sum P_{l,j}^{L} \cdot H_{l,j}} \quad (4.2)$$

and $\bar{v}_{l,j}^{L} = 0.5\,[v_{l,j\,t}^{L} + v_{l,j\,t-1}^{L}]$

where $\Delta \ln L_{l,j}$ indicates the growth rate (measured as long 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. In this approach, which is based on neoclassical theory, the

[12] Aggregate input is measured as a translog index of its individual components. Then the corresponding index is a Törnqvist volume index (see Jorgenson, Gollop and Fraumeni 1987). For all aggregation of quantities, we use the Törnqvist quantity index, which is a discrete time approximation to a Divisia index. This aggregation approach uses annual moving weights based on averages of adjacent points in time. The advantage of the Törnqvist index is that it belongs to the preferred class of superlative indices (Diewert 1976). Moreover, it exactly replicates a translog model which is highly flexible, that is, a model where the aggregate is a linear and quadratic function of the components and time.

implied assumption is that each worker category is paid their marginal productivities[13] and marginal revenues are equal to marginal costs, so the weighting procedure ensures that inputs which have a higher price also have a larger influence in the input index. So, a doubling of number of high-skilled persons worked gets a bigger weight than a doubling of number of low-skilled persons worked. Therefore, the volume growth of labour input as in equation 4.1 can be split into the growth of employment quantity or the number of employed persons and the changes in labour quality index, which is due to the changes in the composition of workers in terms. Let H_j indicate total persons worked in industry 'j' by all types 'l' in period 't', then

$H_{j,t} = \sum_l H_{l,j,t}$,

then we can define the growth in employment as:

$$\Delta \ln H_{j,t} = \Delta \ln \sum H_{l,j,t} \qquad (4.3)$$

and, labour quality growth as:

$$\Delta \ln L_{j,t} - \Delta \ln H_{j,t}$$
$$= \sum \bar{v}_{l,j,t} \Delta \ln H_{l,j,t} - \Delta \ln \sum H_{l,j,t} \qquad (4.4)$$

The first term on the right-hand side indicates the change in labour input and the second term indicates the change in total persons employed in sector 'j'. It can easily be seen that if proportions of each labour type in the labour force change, this will have an impact on the growth of labour input beyond any change in total persons worked. The index of aggregate labour quality thus measures the changes in the composition of labour in the economy.

However, the use of this method becomes more and more data intensive as we introduce more classifications of labour such as gender and age groups. Due to data limitation of sample size becoming very small as we try to estimate persons employed and earnings by industry by all the characteristics,[14] 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[15] has been computed using five education categories[16]

[13] The assumption basically requires perfect competition in labour market, which does not exist in countries like India. It thus restricts the applicability of such a method in situations where there may be widespread monopsony power or bilateral monopoly within an industry.

[14] EU KLEMS faces this problem by restricting the estimation of change in labour quality in some cases to only fifteen aggregate industries and assuming it to be same for sub-industries (Timmer, et al; 2010; p 118).

[15] The index can also be described as "education" index, as only educational characteristic has been used in its computation.

[16] In EU-KLEMS changes in labour quality have been measured by including employment class, gender, age and education (Timmer, op cit.; p.64) and only 3 categories of education, defined as high skilled, medium skilled and low skilled have been taken. This concept of labour quality is refereed as 'labour quality' by Jorgenson. It is recognized (Timmer, et.al.; op cit; p84) that the categorization may lead to biases in the aggregate quality adjustment if employment trends and wage share differ within categories. These education categories generally correspond to- up to Primary education, above Primary to Hr. Secondary education and

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:

i) 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.

ii) 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.

iii) 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, using[17] a Mincer wage equation, which is corrected for sample selection bias by using the Heckman's[18] 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 identification factors used in the first stage are age, gender, marital status, and type of household. 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.

iv) Estimate labour quality index using the equations explained above, and the data compiled in steps i-iii.

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.

---

above Hr Sec or College education. Since due to data limitations, we have measured changes in labour quality only by changes in education profile of labour therefore, it became desirable to have a more detailed classification of education to capture the changes in skill quality of Indian workers.

[17] In PLFS (2017-18), 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.

[18] The details of the function can be obtained from the Stata software and are also provided in Box 1.

Appendix B: Employment Unemployment Survey (EUS) rounds by NSSO and PLFS

NSS Round	Survey Period	MID Period
38	**1/83 to 12/83**	**01-Jul-83**
43	**7/87 to 6/88**	**01-Jan-88**
45	7/89 to 6/90	01-Jan-90
46	7/90 to 6/91	01-Jan-91
47	7/91 to 12/91	01-Oct-91
48	1/92 to 12/92	01-Jul-92
49	1/93 to 6/93	01-Apr-93
50	**7/93 to 6/94**	**01-Jan-94**
51	7/94 to 6/95	01-Jan-95
52	7/95 to 6/96	01-Jan-96
53	1/97 to 12/97	01-Jul-97
54	1/98 to 6/98	01-Apr-98
55	**7/99 to 6/00**	**01-Jan-00**
56	7/00 to 6/01	01-Jan-01
57	7/01 to 6/02	01-Jan-02
58	7/02 to 12/02	01-Oct-02
59	1/03 to 12/03	01-Jul-03
60	01/04 to 6/04	01-Apr-04
61	**7/04 to 6/05**	**01-Jan-05**
64	**7/07 to 6/08**	**01-Jan-08**
66	7/09 to 6/10	01-Jan -10
68	**7/11 to 6/12**	**01-Jan-12**
PLFS (2017-18)	**7/17 to 6/18**	**01-Jan-18**
PLFS (2018-19	**7/18 to 6/19**	**01-Jan-19**

Note: *Highlighted Rounds are major or Quinquennial Rounds of EUS used in the study and rest are annual rounds.*

Appendix C: 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.

Box 1: The Heckman model

To calculate the data on wages for self-employed workers, we estimate a Mincer type regression equation, where wages is a function of worker characteristics (see Mincer, 1974; Heckman, 1976). In doing this, we correct for sample selection bias using Heckman's two-step procedure. In this procedure, we first estimate a selection equation, which helps us understand the selection process. Following the literature, we estimate the selection equation that explains the decision of whether to participate in the labour market, using a Probit model. In our model the dependent variable is a dummy variable that takes the value 1 if a person is working and 0 otherwise, and the independent variables are the relevant individual characteristics of the respondents:

$$z_i^* = w_i \cdot \gamma_i + u_i$$

where $z_i = 1$ if the person i is employed and 0 otherwise. $w_{i,j}$ consists of a set of identification factors, which are age, sex, marital status, and type of household or size of households. Effectively, in this regression, we estimate the effects of individual characteristics on the decision to join the labour market. Since our objective is to estimate the wages of the self-employed, as such, these effects are of least interest to us. But we can use the residual of this Probit regression to obtain information on the effect of the unmeasured characteristics that are not available in the dataset or not captured by the estimated coefficients of the explanatory variables. In the Heckman procedure, these residuals, which are argued to reflect the unmeasured characteristics related to employment, are used to construct a selection bias control factor in a subsequent regression

In the second step, we perform a regression analysis where the dependent variable is earnings of workers and the independent variables are worker characteristics. This regression helps explain wages received by workers, observable only for those for whom the selection equation takes a value of 1.

$y_i = x_i \cdot \beta_i + \varepsilon_i$; observed only if $z_i^* = 1$.

where y_i is the earnings of individual i, $x_{i,j}$'s are a set of j individual characteristics (age, sex, education, location, and marital status) and the industry in which they work. An additional independent variable is added to this equation: the residual (or what we call the inverse of Mill's ratio) from step 1, which captures the unmeasured characteristics.

$(\boldsymbol{u_i}, \boldsymbol{\varepsilon_i})$ ~bivariate normal $[\mathbf{0}, \mathbf{0}, \mathbf{1}, \boldsymbol{\sigma_\varepsilon}$ [0, 0, 1, σε, ρ], where ρ is the correlation between ei and ui. It follows that a standard regression approach would yield biased estimates when ρ ≠ 0. Heckman provides consistent, asymptotically efficient estimates for all the parameters in such a model. In actual estimation, a likelihood ratio test of the independence of these equations testing for ρ = 0 with the corresponding chi-squared statistic is done. This technique helps us overcome the problem of not being able to observe the wage for those who are not employed in the reference period. The regression function has been used to the earnings of casual and regular employees whose wages are observed, and the selection model is used for the unobserved wages of the self-employed persons.

Appendix D: Population (000), WFPR (per thousand) and Persons Employed (000) in Different EUS Rounds

Round↓	Segment→	Rural Males	Rural Females	Urban Males	Urban Females	Total
38th	WFPR	547	340	512	151	
	Population	281288	266637	91217	80445	719587
	Persons Employed	153865	90657	46703	12147	303371
43rd	WFPR	539	323	506	152	
	Population	305500	287900	104400	92800	790600
	Persons Employed	164665	92992	52826	14106	324588
50th	WFPR	553	328	521	155	
	Population	339642	319411	124031	111104	894188
	Persons Employed	187822	104767	64620	17221	374430
55th	WFPR	531	299	518	139	
	Population	374432	353785	145878	131244	1005339
	Persons Employed	198823	105782	75565	18243	398413
61st	WFPR	546	327	549	166	
	Population	400865	379102	164732	148332	1093031
	Persons Employed	218872	123966	90438	24623	457900
68th	WFPR	543.4	248.4	546.4	146.6	
	Population	431722	409834	199785	186101	1227442
	Persons Employed	234600.0	101800.0	109200.0	27300.0	472900
PLFS (2017-18)	WFPR	517	175	530	142	
	Population	459426	435619	227914	214390	1337348
	Persons Employed	237523	76190	120703	30358	464774
PLFS (2018-19	WFPR	521	190	527	145	
	Population	464213	440072	232973	219506	1356763
	Persons Employed	241855	83614	122777	31828	480073

Sources:

1. Employment Unemployment Survey, different issues, and PLFS (2017-18, and 2018-19)
2. Midyear Population Estimates have been obtained from Visaria, (1996) for 1987-88 (Jan); Sundaram, (2007) for 1983 (July), 1993-94 (Jan), 1999-00 (Jan) and 2004-05 (Jan) and from 68th: NSSO Report No 55 Report 554, Appendix- C, page C1-2 for Jan 2012 and from Kannan and Raveendran (2019), EPW, Vol.44, for PLFS- Jan 2018 respectively. Population for 2018-19 has been estimated from the average annual growth of population between 2011 and 2017.

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 2020. All the productivity calculations in India KLEMS use capital services as the 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. The main difference between conventional measures of aggregate capital stock and the measures of capital services used in India KLEMS is that the latter accounts for the differences in marginal productivity between different asset types. While the capital stock measures, when accounted for appropriate depreciation profiles, only consider the differences in productivity between different 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.

5.1 Methodology

To measure capital services, using the Jorgenson approach, we need estimates of capital stock for detailed asset types and the shares of each of these assets in total capital remuneration. Using the Törnqvist approximation to the continuous Divisia index under the assumption of instantaneous adjustability of capital, aggregate capital services growth rate is derived as a weighted growth rate of individual capital assets, the weights being the compensation shares of each asset, *i.e,*

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

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 such that the sum of shares over all capital types add to unity, *i.e.*,

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

where $P^K_{k,t}$ 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 ass 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, $P^K_{k,t}$. Assuming a geometric depreciation rate, δ_k which is constant over time, but different for each asset type, capital stock in asset *k* in and year *t* can be constructed using standard Perpetual Inventory Method (PIM) as: [19]

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

where, $I_{k,t}$ is the real investment in asset type *k* in period t.

The rental price of capital $P^K_{k,t}$ 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):

$$P^K_{k,t} = P^I_{k,t-1} \cdot i_t + P^I_{k,t} \cdot \delta_k - (P^I_{k,t} - P^I_{k,t-1}) \tag{5.4}$$

where $P^I_{k,t}$ 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. [20] The rate of return *i* in equation (5.4) can be measured either as an internal or

[19] 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: $P^I S_t = \sum_k P^I_{k,t} S_{k,t}$, where $P^I_{k,t}$is the investment deflator for asset k, S_(k,t) is the capital stock in asset k.

[20] 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: $(1+i_t)P^I_{k,t-1} = P^K_{k,t} + (1-\delta_k)P^I_{k,t}$, 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: $p^K_{k,t} = p^I_{k,t-1} i_t + \delta_k p^I_{k,t} - [p^I_{k,t} - p^I_{k,t-1}]$.

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).

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).[21] We exploit multiple 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 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 KLEMS[22], 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.

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 also 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$\left(KQ_t\right)$, i.e. $KQ_t = KQ_{t-1} \times \mathrm{e}^{(\Delta \ln K_t - \Delta \ln S_t)}$ for t>base year, and $KQ_t = 1$ for t=base year. In the India KLEMS database, the capital composition growth rates are indexed to the first year, 1980.

---

[21] 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 Erumban and Das (2015).

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

5.2 Data and sources

(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.[23] 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. However, there were slight differences between the industry groups available in the current release of the NAS data and the previous ones (see Table 5.1), which required some matching of sectors before combining the two series. Table 5.2 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.

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

Table 5. 1: Industries for which GFCF data by assets are available in NAS

NAS 2015 (1950-2012)	NAS 2016 (2011-2016)*	NAS annual series (since 2016)	India KLEMS sector	ISIC 3.1
Agriculture	Crops			
Forestry & logging	Forestry & logging	crops + livestock	Agriculture, hunting, forestry & fishing	
Fishing	Fishing	fishing and aquaculture		AtB
Mining & Quarrying	Mining & Quarrying	Mining & Quarrying	Mining & quarrying	C
Reg. Manufg.	Reg. Manufg.	Manufacturing (Public sector, Corporate sector & household sector)		
Unreg. Manufg.	Unreg. Manufg.		Manufacturing	D
Electricity	Electricity			
Gas	Gas			
Water Supply	Water Supply	Electricity, gas, water supply and other utility services	Electricity, gas & water supply	
N.A.	Remaining utilities			E
Construction	Construction	Construction	Construction	F
Trade	Trade	Trade & repair services	Trade	G
Hotels & Restaurants	Hotels & Restaurants	Hotels & restaurants	Hotels & restaurants	H
Railways	Railways	Railways		
Other public transport	N.A	Transport, storage, communication & services related to broadcasting		
N.A.	Road Transport	Road Transport		
N.A.	Water Transport	Water Transport		
N.A.	Air Transport	Air transport		
N.A.	Services incidental to transport			
Other transport, private	N.A			
Storage	Storage	Storage	Transport & Storage	60t63
Communication	Communication	Communication & services related to broadcasting	Post & telecommunication	64
Banking & Insurance	Financial Services	Financial services	Financial services	J
Real Estate	Real estate	Real estate, ownership of dwelling and professional services		
N.A.	Ownership of dwellings		Other services	70+O+P
Public admin. & defence	Public admin. & defence	Public admin. & defence	Public administration & defense; compulsory social security	L
Other Services	N.A.	Other services	Other services (incl. education & health)	70+M to P
N.A.	Education		Education	M
N.A.	Health		Health & social work	N
N.A.	Other services		Other services	70+O+P

Note: *Some of the India KLEMS sectors listed in the last two columns are further split into sub sectors, using additional data from ASI and NSSO. * Consistent with 2011-2012 base NAS data*

Table 5. 2: 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[24]	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**

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.*

Source: CSO, NAS Different Issues

Total investment in each asset category is calculated as the sum of private and public sector investment in each asset. Investment in transport equipment is not available separately for private sector. We tried several approaches to impute the private sector transport equipment

[24] 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.

data. The first was to use the share of transport equipment in non-departmental enterprises. More specifically, we apply the non-departmental enterprise transport equipment to machinery & equipment (including transport equipment) ratio to machinery & equipment in private sector for each industry to obtain industry wise transport equipment for private sector. We take non-departmental enterprise only, rather than the entire public sector, as it may be more realistic as it consists of public sector companies and statutory corporations, excluding administrative sector. However, the sum of industry estimates generated by this approach was not consistent with the reported aggregate private sector transport equipment. Therefore, we take a second step here, which is to use the industry distribution from this series and apply it to the published total private sector transport equipment data. The estimated transport equipment is then subtracted from each industry's total machinery & transport equipment data, to obtain machinery in private sector as a residual. However, this approach generates many negative numbers in transport equipment in private sector, particularly in transport services. Therefore, we follow a third approach, which is what we finally use in the database. We first distribute 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.

However, since 20216-2017, as we did not obtain detailed asset/industry-wise data on capital formation from the CSO, we relied solely on publicly available data from the National Accounts. The publicly available data contains more detailed information now, compared to what it used to be in the past. Therefore, the reliance on the NAS public data did not create too many challenges. However, there were two outstanding issues. The first was regarding the imputation of transport equipment. Data on transport equipment was not separately available in this data. Therefore, 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)*).

The second issue was related to the split of the manufacturing sector into registered and unregistered manufacturing. Since NAS provides the investment data for the aggregate manufacturing sector, we use information from the Annual Survey of Industries (ASI) and NSSO unorganized surveys to distribute the aggregate GFCF across industries (see section b). However, NAS does not provide the split between registered and unregistered manufacturing since 2016, making it challenging to use ASI and NSSO data to split the aggregate manufacturing. Nevertheless, NAS provides a division between the corporate sector, public sector, and household sector. As a quick fix, 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 release.

As discussed at the beginning of this section, 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,...,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:

$GFCF_{j,t}^{2012} = \sum_{i=1}^{n} GFCF_{i,t}^{2012}$; 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:

$GFCF_{i,t}^{2012} = \left[\frac{GFCF_{i,t}^{2004}}{GFCF_{i,t+1}^{2004}}\right] GFCF_{i,t+1}^{2012}$ 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:

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

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

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.

(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 capital/labour ratio in the higher aggregate industry. However, the final data used are based on value added shares, as a sensitivity analysis did not show a 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. However, the published data does not provide any asset wise investment information; it consists of only the aggregate capital formation or the book value of fixed capital. Most studies in the past have measured gross investment as the difference between book value of asset in period t and in period t-1 and add depreciation in period t to that. This approach has the deficiency of comparing two different samples reported in two different years, where the number of firms/factories might be different. In particular, while using this approach at industry level, for detailed asset categories, it might generate massive negative investment.

We follow an alternative approach, following ASI’s definition of gross fixed capital formation (GFCF). 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 method. Table 5.3 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.

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 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.

Table 5. 3: 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*

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 tables. 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. Asset classification in NSSO has changed over various rounds, and therefore, we have tried to match these with our classification as shown in Table 5.4. 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.

Table 5. 4: **Asset Categories in NSSO Rounds**

	NSSO sectors				
45th Round (1989-90)	**51st Round (1994-95)**	**56th Round (2000-01)**	**62nd Round (2005-06)**	**67th & 73rd round (2010-11 & 2015-16)**	**India KLEMS Asset Types**
Land	Land			Land	Excluded
	Building			Building	Construction
		Land & Buildings	Land & Buildings		Construction (land is excluded)
	Other construction				Construction
Building & other construction					Construction
Plant & machinery	Plant & machinery	Plant & machinery	Plant & machinery	Plant & Machinery	Machinery & Equipment
Transport Equipment	Transport Equipment	Transport equipment	Transport equipment	Transport equipment	Transport Equipment
	Tools				Machinery & Equipment
	Other fixed assets				Machinery & Equipment
Tools & other fixed assets		Tools & other fixed assets	Tools & other fixed assets	Tools & other fixed assets	Machinery & Equipment
			Software & hardware		Machinery & Equipment*
				Software & databases	Machinery & Equipment*
				Information & Communication equipment	Machinery & Equipment*

Note: *For 56th and 62nd rounds, land is separated from land & buildings using land/land & building ratio from 51st round.*
** 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: *NSSO Rounds*

(c) Investment Prices by Asset Types

In order to compute asset wise capital stock using PIM (equation 5.3) and rental price (equation 5.4), 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

As is evident from equations 5.1 to 5.4, our estimates of capital input require time-series data on asset wise capital stock. Capital stock has been constructed using perpetual inventory method (PIM), where the capital stock (S) is defined as a weighted sum of past investments with weights given by the relative efficiencies of capital goods at different ages, which requires data on current investment by asset types, investment prices by asset types and depreciation rate. Also, for the practical 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.[25] 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.[26] 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 5.5 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.

---

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

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

Table 5. 5: 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 equation 5.4). 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 *ex post*) rate of return, or as an external (*ex ante)* rate of return.[27] This issue will be addressed in the further revisions of the data. 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 India[28]. Therefore, we use a real rate, which is net of capital gain. Hence, the capital gain component in equation (5.4) is excluded while estimating rental price using external rate of return, obtaining

$$P^K_{k,t} = P^I_{k,t-1} \cdot r_t + P^I_{k,t} \cdot \delta_k \qquad (5.5)$$

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

[27] 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.

[28] Reserve Bank of India, Handbook of Indian Statistics, Annual volumes.

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 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.

****

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). This breakdown of intermediate inputs can be used for extending the growth accounting exercises, but also convey interesting information about changing pattern in intermediate consumption (see *e.g.*, JHS 2005, chapter 4)

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. Every commodity is accounted for, whether produced by a domestic source or imported, and every use is noted, whether purchased by an industry or by a final demand element. All payments to factors of production *i.e.*, labour and capital is accounted for so that all income elements of GDP are included. The methodology of constructing time series on energy, material and services inputs for the European economy has been elucidated in Timmer *et al.* (2010, Chapter 3). 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**

Intermediate Inputs are broken down into energy, material and services, based on input output transaction tables using a standard NIC product classification. 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**.

The key building block for constructing time series on Intermediate Inputs at current prices, as explained in Jorgenson *et al.* (2005, Chapter 4), is the input-output transaction tables, that is, the inter industry transaction tables that provide a description of which industries produce each product and which industries use them. The input-output table gives the inter-industry transactions in value terms at factor cost presented in the form of commodity x industry matrix where the columns represent the industries and the rows as group of commodities, which are the principal products of the corresponding industries. Each row of the matrix shows in the relevant columns, the deliveries of the total output of the commodities to the different industries for intermediate consumption and final use. The entries read down industry columns give the commodity inputs of raw-materials and services, which are used to produce outputs of particular industries. The column entries at the bottom of the table give net indirect taxes (NIT) (indirect taxes – subsidies) on the inputs and the primary inputs (income from use of labour and capital), *i.e.*, Gross Value Added (GVA). As the IOTT is in the form of commodity x industry matrix, the row totals do not tally with the column totals. The difference between each column and the corresponding row totals is due to the inclusion of the secondary products, which appear particularly in the case of manufacturing industries. This is so because by-products are also manufactured by industries in addition to their main products. Thus, while determining the entries in the rows, a by-product of an industry is transferred to the sector (commodity row), whose principal product is the same as the by-product under reference. The columns, however, show the total of principal products and by-products of each industry. All the entries in the IOTT are at **factor cost**, *i.e.*, excluding trade and transport charges and NIT.

The Input Flow Matrix at factor cost, published by CSO, for 1978 is a 60 x 60 matrix. The absorption matrices for 1983, 1989, 1993 and 1998 have 115 sectors. However, a detailed 130 sector absorption (commodity x industry) matrix for the Indian economy has been published from 2003-04 onwards. The scheme of sector classification adopted in IOTT 2015-16[29], IOTT 2007-08 and IOTT 2003-04 *vis-à-vis*, IOTT 1983-84, IOTT 1989-90, IOTT 1993-94 and 1998-99 has undergone significant change with the disaggregation of some of the sectors, which have become significant in early 2000s.

1. In the agriculture and allied activities, four sectors, namely, 'oil seeds', 'fruits', 'vegetables' and 'poultry & eggs' have been introduced as separate sectors.
2. In the mining sectors, the sector 'crude petroleum, natural gas' has been disaggregated into two separate sectors as 'natural gas' and 'crude petroleum'.
3. The sector 'office computing machines' in the earlier IOTT has been eliminated.
4. The sector 'miscellaneous manufacturing' in the earlier IOTT has been disaggregated into 'medical, precision& optical instruments', 'gems & jewelry', 'aircraft & spacecraft' and 'miscellaneous manufacturing'.
5. Similarly, 'other transport services' in the earlier IOTTs are further disaggregated into 'land transport including via 'pipelines', 'water transport', 'air transport' and 'supporting and auxiliary transport activities'.
6. Similarly, sector 'gas' has been regrouped in the present IOTT by merging its gobar gas component in the 'other livestock product' and LPG component in 'petroleum products'.
7. Another disaggregation is done in the 'other services' sector by reorganizing this into separate sectors as 'business services', 'computer related services', 'legal services', 'real estate service activities', 'renting of machinery & equipment', 'other community, social & personal services' and 'other services'.

6.2 Implementation Procedure

The methodology for computation of Intermediate Input Series for 27 Industries from 1980-2018 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. Concordance table (1998 and 2003) has been attached at the end (Refer Appendix Table E below).

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

[29] IOTT 2015-16 was published by Brookings India.

industries from 1980 to 2018. 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 2018) 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 2018.
- Thus, using IOTT, the intermediate input vector has been projected for 27 study industries from 1980 to 2018.

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. This has been done because the relevant proportions in SUT 2015-16 is believed to be making a more correct assessment.

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. For example, coal consumed in the Coke industry is taken primarily as material rather than energy. From IOTT it is difficult to estimate how much of coal consumption is for energy purposes and how much as material input, where ASI provides details information about energy consumption by energy inputs like: coal, electricity, petroleum and others including natural gas.

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. Two examples have been given below for the Construction sector:

Adjustment A: For IOTT Benchmark Year 1998

Source	Output/Input	Values (in Rs Crore)	
National Account Statistics	GVO	241873	
	GVA	88784	
	Total Input from NAS (GAP)	153089	
For Benchmark IOTT year: Absolute value of Inputs from Input Output Transaction Table are available	Material	73132	
	Energy	4308	
	Service	37399	
	Total Input from IOTT	114839	
Ratio of Total Input from NAS to Total Input from IOTT	Total Input from NAS/ Total Input from IOTT = 1.33		
Adjusted Input reported	Material	= 1.33 x 73132	97490
	Energy	= 1.33 x 4308	5743
	Service	= 1.33 x 37399	49855
	Total Input	= 1.33 x 114839	153089

Adjustment B: For Non-Benchmark Year 1999

Source	Output/Input	Values (in Rs Crore)	
National Account Statistics	GVO	279464	
	GVA	102007	
	Gap	177457	
For Non-Benchmark IOTT year: Interpolated Proportions of Inputs from Input Output Transaction Table have been used	Material	0.64	
	Energy	0.04	
	Service	0.32	
	Total Input	1	
Adjusted Input reported	Material	= .64 x 177457	113081
	Energy	= .04 x 177457	7300
	Service	= .32 x 177457	57076
	Total Input	= 1 x 177457	177457

In the examples above, we can see that the total value of intermediate input exactly matches the gap between Gross output and Value added. Similar adjustments have been made to construct the input series for all the other 27 study industries from 1980-2018. For every benchmark year where IOTT is available adjustment 'A' has been done. For every non-benchmark year, where IOTT is not available and an intermediate input vector has been projected, adjustment 'B' has been done to generate the comprehensive time series of Intermediate inputs consistent with the official National Accounts.

Steps 1-4: This gives a time series of Material, Energy, and Service Inputs for 27 study Industries from 1980 to 2018 at current prices.

Steps 5 and 6 below, explain the methodology for computation of Intermediate Input Series for 27 study Industries from 1980-2018 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 2018 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 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 2018. 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 1980-2018, 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.

6.3 Outstanding Issues

Firstly, input output transaction tables are generally available on a five-year interval period and this necessitates interpolation and assumption of constant shares in some cases, to construct the entire time series of EMS from 1980 to 2018.

Secondly, unlike studies using detailed survey data, we have to assume that all buyers pay the same price for each commodity because there is no information about price divergences.

Thirdly, the estimated time series of Intermediate Inputs at constant price will not be consistent with the intermediate input of NAS at constant price *i.e.*, the deflated value of intermediate input cannot match the gap between value added and gross output at constant price from NAS. This is because; NAS uses a single deflation method to estimate Gross Value Added and Gross Output at constant price.

Appendix E: Concordance table of study industries and IOTT industries

KLEMS Industry No.	Industry Description	Sector No. IOTT 2003	Sector description	Sector No. IOTT 1998	Sector
1	Agriculture, Hunting, Forestry and Fishing	1	Paddy	1	Paddy
		2	Wheat	2	Wheat
		3	Jowar	3	Jowar
		4	Bajra	4	Bajra
		5	Maize	5	Maize
		6	Gram	6	Gram
		7	Pulses	7	Pulses
		8	Sugarcane	8	Sugarcane
		9	Groundnut	9	Groundnut
		10	Jute	10	Jute
		11	Other Oil Seed	11	Cotton
		12	Cotton	12	Tea
		13	Tea	13	Coffee
		14	Coffee	14	Rubber
		15	Rubber	15	Coconut
		16	Coconut	16	Tobacco
		17	Tobacco	17	Other crops
		18	Fruits		
		19	Vegetables		
		20	Other crops		
		21	Milk and milk products	18	Milk and milk products
		22	Animal services(agricultural)	19	Animal services(agricultural)
		23	Poultry & eggs		
		24	Other livestock products	20	Other livestock products
		25	Forestry and logging	21	Forestry and logging
		26	Fishing	22	Fishing
2	Mining & Quarrying	27	Coal and lignite	23	Coal and lignite

		28	natural gas	24	Crude petroleum, natural gas
		29	Crude petroleum		
		30	Iron ore	25	Iron ore
		31	Manganese ore	26	Manganese ore
		32	Bauxite	27	Bauxite
		33	Copper ore	28	Copper ore
		34	Other metallic minerals	29	Other metallic minerals
		35	Lime stone	30	Lime stone
		36	Mica	H	Mica
		37	Other non-metallic minerals	32	Other non-metallic minerals
3	Food Products, Beverages and Tobacco	38	Sugar	33	Sugar
		39	Khandsari, boora	34	Khandsari, boora
		40	Hydrogenated oil(vanaspati)	35	Hydrogenated oil(vanaspati)
		41	Edible oils other than vanaspati	36	Edible oils other than vanaspati
		42	Tea and coffee processing	37	Tea and coffee processing
		43	Miscellaneous food products	38	Miscellaneous food products
		44	Beverages	39	Beverages
		45	Tobacco products	40	Tobacco products
4	Textiles, Textile Products, Leather and Footwear	46	Khadi, cotton textiles(handlooms)	41	Khadi, cotton textiles (handlooms)
		47	Cotton textiles	42	Cotton textiles
		48	Woolen textiles	43	Woolen textiles
		49	Silk textiles	44	Silk textiles
		50	Art silk, synthetic fiber textiles	45	Art silk, synthetic fiber textiles
		51	Jute, hemp, mesta textiles	46	Jute, hemp, mesta textiles
		52	Carpet weaving	47	Carpet weaving
		53	Readymade garments	48	Readymade garments
		54	Miscellaneous textile products	49	Miscellaneous textile products

		59	Leather footwear	54	Leather footwear
		60	Leather and leather products	55	Leather and leather products
5	Wood and Products of Wood	56	Wood and wood products	51	Wood and wood products
6	Pulp, Paper, Paper Products, Printing and Publishing	57	Paper, paper prods. & newsprint	52	Paper, paper prods. & newsprint
		58	Printing and publishing	53	Printing and publishing
7	Coke, Refined Petroleum Products and Nuclear Fuel	63	Petroleum products	58	Petroleum products
		64	Coal tar products	59	Coal tar products
8	Chemicals and Chemical Products	65	Inorganic heavy chemicals	60	Inorganic heavy chemicals
		66	Organic heavy chemicals	61	Organic heavy chemicals
		67	Fertilizers	62	Fertilizers
		68	Pesticides	63	Pesticides
		69	Paints, varnishes and lacquers	64	Paints, varnishes and lacquers
		70	Drugs and medicines	65	Drugs and medicines
		71	Soaps, cosmetics & glycerin	66	Soaps, cosmetics & glycerin
		72	Synthetic fibers, resin	67	Synthetic fibers, resin
		73	Other chemicals	68	Other chemicals
9	Rubber and Plastic Products	61	Rubber products	56	Rubber products
		62	Plastic products	57	Plastic products
10	Other Non-Metallic Mineral Products	74	Structural clay products	69	Structural clay products
		75	Cement	70	Cement
		76	Other non-metallic mineral prods.	71	Other non-metallic mineral prods.
11	Basic Metals and Fabricated Metal Products	77	Iron, steel and ferro alloys	72	Iron, steel and ferro alloys

		78	Iron and steel casting & forging	73	Iron and steel casting & forging
		79	Iron and steel foundries	74	Iron and steel foundries
		80	Non-ferrous basic metals	75	Non-ferrous basic metals
		81	Hand tools, hardware	76	Hand tools, hardware
		82	Miscellaneous metal products	77	Miscellaneous metal products
12	Machinery, nec	83	Tractors and agri. implements	78	Tractors and agri. implements
		84	Industrial machinery (F & T)	79	Industrial machinery (F & T)
		85	Industrial machinery(others)	80	Industrial machinery(others)
		86	Machine tools	81	Machine tools
		87	Other non-electrical machinery	82	Office computing machines
				83	Other non-electrical machinery
13	Electrical and Optical Equipment	88	Electrical industrial Machinery	84	Electrical industrial Machinery
		89	Electrical wires & cables	85	Electrical wires & cables
		90	Batteries	86	Batteries
		91	Electrical appliances	87	Electrical appliances
		92	Communication equipments	88	Communication equipments
		93	Other electrical Machinery	89	Other electrical Machinery
		94	Electronic equipments(incl.TV)	90	Electronic equipments(incl.TV)
		101	Watches and clocks	97	Watches and clocks
		102	medical, precision and optical instruments		
14	Transport Equipment	95	Ships and boats	91	Ships and boats

		96	Rail equipments	92	Rail equipments
		97	Motor vehicles	93	Motor vehicles
		98	Motor cycles and scooters	94	Motor cycles and scooters
		99	Bicycles, cycle-rickshaw	95	Bicycles, cycle-rickshaw
		100	Other transport equipments	96	Other transport equipments
		104	Air craft & space craft		
15	Manufacturing, nec; recycling	55	Furniture and fixtures-wooden	50	Furniture and fixtures-wooden
		103	gems & jewellery		
		105	Miscellaneous manufacturing	98	Miscellaneous manufacturing
16	Electricity, Gas and Water Supply	107	Electricity	100	Electricity
		108	Water supply	101	Gas
				102	Water supply
17	Construction	106	Construction	99	Construction
18	Trade	116	Trade	107	Trade
19	Hotels and Restaurants	117	Hotels and restaurants	108	Hotels and restaurants
20	Transport and Storage	109	Railway transport services	103	Railway transport services
		110	Land transport including via pipeline	104	Other transport services
		111	Water Transport	105	Storage and warehousing
		112	Air Transport		
		113	Supporting and Auxilary Transport Activities		
		114	Storage and warehousing		
21	Post and Telecommunication	115	Communication	106	Communication
22	Financial Services	118	Banking	109	Banking
		119	Insurance	110	Insurance
23	Business Service	123	Business services	114	Other services
		124	Computer and related activities		
		125	Legal services		
			Renting of		

		127	machinery and equipment		
24	Public Administration and Defense	130	Public administration	115	Public administration
25	Education	121	Education and research	112	Education and research
26	Health and Social Work	122	Medical and health	113	Medical and health
27	Other Services	120	Ownership of dwellings	111	Ownership of dwellings
				114	Other services
		126	Real estate activities		
		128	Other communication, social and personal services		
		129	Other services		

Source: *Input Output Transaction Table 1998, 2003.*

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, OS 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 Intermediation
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
	27	Other 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. In the Back Series with base 2011-12, NAS updated the estimates of GVA for the years prior to 2011-12, but did not provides the updated estimates of CE, OS and MI. Therefore, for 1980-81 to 2011-12 period, we proportionally adjusted the CE, OS and MI series corresponding to the updated GVA estimates.

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. In certain cases, the estimates of CE, OS and MI for a particular NAS sector have been distributed across constituent study industries proportionately in accordance with the gross value added in those industries. Estimate of factor incomes for ‘other services’ in NAS, for instance, 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:

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

In this equation, SL^i_t is the labour income share in industry i in year t, CE^i_t is compensation of employees in industry i in year t, MI^i_t is mixed income of the self-employed persons in industry i in year t, and GVA^i_t 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

ηobtained 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 2017-18 and 2018-19.

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 annul 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.[30] 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.

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.

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

(c) Having obtained an estimate of η, equation 6.1 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 Intermediation
	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:

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

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 outputs.[31] 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.

In order to estimate the production function at constant prices, the industry level price data is used for output, capital, labour and intermediate inputs, *e.g.*, PY_j; PL_j, PK_j; PE_j. PM_j, PS_j. We assume that all industries face the same input price. 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.

[31] 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 equipments 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).

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.

$$\boldsymbol{P_{Y,j}Y_j = P_{K,j}K_j + P_{L,j}L_j + P_{E,j}E_j + P_{M,j}M_j + P_{S,j}S_j} \qquad \textbf{(8.2)}$$

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

$$\boldsymbol{v_{T,j} \equiv \Delta lnY_j - \bar{v}_{K,j}\Delta lnK_j - \bar{v}_{L,j}\Delta lnL_j - \bar{v}_{E,j}\Delta lnE_j - \bar{v}_{M,j}\Delta lnM_j - \bar{v}_{S,j}\Delta lnS_j} \qquad \textbf{(8.3)}$$

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

We define the value share of each input as:

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

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

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

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

The assumption of constant returns to scale implies $\boldsymbol{\sum_i v_{i,j} = 1}$ 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

$$\boldsymbol{\Delta lnY_j = \bar{v}_{K,j}\Delta lnK_j + \bar{v}_{L,j}\Delta lnL_j + \bar{v}_{E,j}\Delta lnE_j + \bar{v}_{M,j}\Delta lnM_j + \bar{v}_{S,j}\Delta lnS_j + v_{T,j}} \qquad \textbf{(8.6)}$$

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

$$\boldsymbol{\Delta lnY_j = \bar{v}_{K,j}\,(\Delta lnZ_j + \Delta lnQ_{K,j}) + \bar{v}_{L,j}\left(\Delta lnH_j + \Delta lnQ_{L,j}\right) + \bar{v}_{E,j}\Delta lnE_j + \bar{v}_{M,j}\Delta lnM_j + \bar{v}_{S,j}\Delta lnS_j + v_{T,j}} \qquad \textbf{(8.7)}$$

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.

8.2 Methodology for Aggregation across Industries

In the analysis presented in the above section, gross value of output (GVO) is taken as the measure of output of an industry and TFP growth was estimated using the gross output function framework in which capital, labour, materials, energy and services are taken as five inputs. The Törnqvist index is applied to estimate TFP growth for each year during 1980-81 to 2018-19, which yielded an index of TFP for each industry for that period, permitting estimation of trend growth rate in TFP for the period under study (1980-81 to 2018-19).

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. It becomes necessary therefore to consider appropriate methods of aggregation across industries consistent with the gross output function specification of technology at the individual industry level. It is needless to say that one has to use the concept of value added to define an appropriate measure of output at the economy or broad sector level, but even here there are important issues of aggregation, *i.e.*, how value added in different industries should be combined. A very useful discussion of the issues involved is found in Jorgenson *et al.* (2005).

There are three approaches to estimating TFP growth at an aggregate level. This is discussed in detail in Jorgenson *et al.* (2005). The first approach is called aggregate production function approach. This approach assumes the existence of an aggregate production function (say, at the level of the economy or at the level of manufacturing or services sector). An essential condition is that the production function be separable in primary inputs. Let the production function for a particular industry be defined as:

$$Y = f(X, K, L, T), \qquad \textbf{(8.8)}$$

where Y denotes gross output, X intermediate input (in turn a combination of materials, energy and services), K capital input, L labour input and T time (representing technology). Then, the concept of value added requires the existence of a value-added function which in turn requires that the above production function be separable in K, L and T, *i.e.*, the production function should take the following form:

$$Y = g(X, V(K, L, T)). \qquad \textbf{(8.9)}$$

In this equation, V (.) is the value-added function. V denotes value added.

Besides the condition of separability of the production function described above, a number of highly restrictive assumptions have to be made for the aggregate production function approach. These include: (a) the value-added function is the same across all industries up to a scalar multiple, (b) the functions that aggregate heterogeneous types of labour and capital must be identical in all industries, and (c) each specific type of capital and labour receives the same price in all industries. With these assumptions made, real value added at the aggregate level becomes a simple addition of real value added in individual industries. In notation,

$$V = \sum_i V_i \qquad \textbf{(8.10)}$$

where V_i is value added in the i^{th} industry and V is the aggregate value added. Given aggregate value added and somewhat similarly defined aggregate capital and labour input, TFP growth at the aggregate level can be computed.

The second approach to aggregation is the aggregate production possibility frontier approach. It also needs the separability condition described in equations 8.8 and 8.9 above, as in the case of the aggregate production function approach. The main difference between the aggregate production function approach and the aggregate production possibility frontier approach is that the latter relaxes the assumption that all industries must face the same value added function. Thus, the price of value added is no longer assumed to be the same across industries. The implication is that the aggregate value added from the aggregate production possibility frontier is given by the Törnqvist index of the industry value added as:

$$\Delta \ln V = \sum_i \bar{w}_i \, \Delta ln V_i \qquad \textbf{(8.11)}$$

In this equation, w_i is the share of industry *i* in aggregate value added in nominal terms. Thus, defining $P_{V,i}$ as the price of value added in industry i and V_i as the real value added in industry i, the share in question may be defined as:

$$w_i = \frac{P_{V,i}V_i}{\sum_j P_{V,j}V_j} \;, \qquad \textbf{(8.12)}$$

and the two-period average is defined as:

$$\bar{w}_i = 0.5 * (w_{i,t} + w_{i,t-1}) \qquad \textbf{(8.13)}$$

Since each industry is subject to a production function separable in capital, labour and technology, as defined in equation 8.9, real gross output, real value of intermediate inputs and real value added of industry i in a particular year t should satisfy the following relationship:

$$\Delta \ln Y_i = \bar{u}_{V,i} \Delta ln\, V_i + \bar{u}_{X,i} \Delta ln X_i \qquad \textbf{(8.14)}$$

In this equation, u_V and u_X are the shares of value added and intermediate inputs in the gross value of output in nominal terms. Given data on growth in gross output and growth in intermediate input of a particular industry, the above equation yields the growth in real value added of that industry.

TFP growth from the aggregate production possibility frontier may be defined as follows:

$$v_T = \Delta lnV - \bar{v}_K \Delta lnK - \bar{v}_L \Delta lnL \quad \textbf{(8.15)}$$

In this equation K and L are aggregate capital input and labour input respectively, and v_K and v_L are value shares (or income shares) of capital and labour respectively. If one maintains the assumption that each specific type of capital and labour input has the same price in all industries, then each type of capital (labour) can be summed across industries and then a Törnqvist index can be constructed to yield aggregate capital (labour) input. Alternatively, different types of capital input can be combined using a Törnqvist index into total capital input used in an industry and then industry level growth in capital input can be combined with the help of a Törnqvist index to obtain growth in capital input at the aggregate level. In a similar manner, the growth in aggregate labour input can be computed.

The third approach to measuring the TFP index at an aggregate level is to apply direct aggregation to industry level estimates. This maintains the industry level production accounts as the fundamental building block and begins with industry level sources of growth. Of the three approaches described here, this is the least restrictive in terms of the assumptions involved. The following equation expresses the relationship between aggregate value added growth and the growth in capital and labour inputs and TFP in individual industries:

$$\Delta \ln V = \sum_i \bar{w}_i \frac{\bar{v}_{K,i}}{\bar{v}_{V,i}} \Delta \ln K_i + \bar{w}_i \frac{\bar{v}_{l,i}}{\bar{v}_{V,i}} \Delta lnL_i + \bar{w}_i \frac{1}{\bar{v}_{V,i}} v_{T,i} \quad \textbf{(8.16)}$$

In this equation, v_K is the value share of capital in gross output, v_L is the value share of labour in gross output, v_V is the share of value added in gross output, and w_i is the share of industry i in aggregate value added in nominal terms. The bars indicate that the average of two periods, t and t-1, are to be taken. The last term in equation 8.16 is:

$$\sum_i \bar{w}_i \frac{1}{\bar{v}_{V,i}} v_{T,i} = v_T^D \quad \textbf{(8.17)}$$

This is the Domar aggregation of TFP growth rates at individual industry level. Note that it is weighted average of industry level TFP growth rates. However, the weights add up to more than one.

8.3 Implementation Procedure

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 2018. 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 framework (i.e. equation 8.7). 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.

Turning now to the aggregate level estimates of TFPG, the three alternate approaches to aggregation described in Section 8.2 above, drawing on Jorgenson *et al.* (2005), are applicable to economy level aggregation, but these are not readily applicable to intermediate level of aggregation, for instance, aggregation to the level of the manufacturing sector. In the India KLEMS dataset, aggregate level real GVA growth, labour input growth, capital input growth and TFP growth based on the value added function framework have been provided for the aggregate economy and for the manufacturing sector and the services sector. Of the three approaches to aggregation discussed in Section 8.2 above, the third approach which involves application of Domar weights is relatively superior as it makes least restrictive assumptions among the three approaches. However, the manner of constructing applicable Domar weights for obtaining the TFPG estimates for the manufacturing sector aggregate and the services sector aggregate differ from those applicable for making economy level TFPG estimates. In order to ensure that the estimates of TFPG obtained for the manufacturing sector and the services sector (and provided in the India KLEMS dataset) are comparable to those for the aggregate economy, a rather simple approach has been taken. This approach is based on application of the Törnqvist index for the purpose of computing the growth rates in real GVA, labour input and capital input and deriving the TFP growth rate accordingly, based on the value added function framework. Taking this approach, computations are made for the manufacturing sector and the services sector, and for the aggregate economy. The simple approach that has been taken does not have as good a conceptual basis as the three approaches described in the previous section particularly the approach of direct aggregation, but it has the advantage of easy implementation and comparability of estimates. The methodology that has been applied to obtain TFPG for the manufacturing sector and services sector and the economy is spelt out further in what follows.

In the India KLEMS database, we only provide simple Tornquvist aggregates of TFP growth rates for broad sectors and the aggregate economy. These aggregates are obtained using a Tornqvist aggregate of growth rates of real value added, capital input, and labour input. Suppose the aggregate sector J consists of n KLEMS industries (j=1:n). We obtain the output and inputs aggregates for sector J as:

$$\Delta lnV_J = \sum_j \bar{v}_j \, \Delta lnV_j \qquad \textbf{(8.18)}$$

$$\Delta lnK_J = \sum_j \overline{w}_j^K \, \Delta lnK_j \qquad \textbf{(8.19)}$$

$$\Delta lnL_J = \sum_j \overline{w}_j^L \, \Delta lnL_j \qquad \textbf{(8.20)}$$

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

$$v_j = \frac{P_{V,j}V_j}{\sum_j P_{V,j}V_j}; w_j^K = \frac{P_{K,j}S_j}{\sum_j P_{K,j}S_j}; and\ w_j^L = \frac{P_{L,j}H_j}{\sum_j P_{L,j}H_j} \quad \textbf{(8.21)}$$

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 and input growth rates obtained above are then used with the growth accounting equation to obtain the sectoral TFP growth rates (see equation 8.3; an equation similar to that equation is used in which only value added, labour input and capital input are considered along with the income shares of labour and capital in gross value added).