

# An Employment Data Strategy for India<sup>§</sup>

**ABSTRACT** Following the recommendations of the Task Force set up by the Government of India in 2017 to revamp the employment data architecture, India's employment statistics have undergone an overhaul. Two significant changes have been the replacement of NSSO's quinquennial Employment and Unemployment Surveys with the annual Periodic Labour Force Survey (PLFS) and the introduction of monthly payroll data. Given the dualistic nature of India's labor markets and the dominance of low-wage and low-productivity informal jobs, this exercise will serve a limited purpose. Against this backdrop, we examine the inadequacies of the existing data architecture, identify the gaps in data collection, and make recommendations for generating more relevant and comprehensive labor market data. We argue that it is imperative to continue with NSSO's quinquennial household survey and supplement it with the annual PLFS. The collection of data in household surveys needs to be based on all three different employment approaches, usual status, current weekly status, and current daily status. None of these can be dispensed with. We propose constructing measures of labor underutilization and informal employment to help us understand the nature and enormity of the employment challenge. We argue that there is a need to rethink the existing classification of employment status categories and to ask more probing questions about the nature of employment and activities undertaken by workers. We also highlight the importance of establishing a Business Register, as enterprise surveys face the problem of an incomplete frame. Instead of simply conducting new surveys, we believe it would be more prudent to strengthen the existing data machinery and ensure that state governments are equal partners in this exercise.

**Keywords:** *Labor Force, Employment, Employment Data, Household Surveys, Enterprise Surveys*

**JEL Classification:** *J21, J29, J40, J46, J60*

\* *radhicka.kapoor@gmail.com*

§ The author is extremely grateful to Ajit Ghose, G. C. Manna, P. C. Mohanan, Rinku Murgai, R. Nagaraj, and participants at the India Policy Forum for their insightful comments and suggestions. Ashulipi Singhal provided excellent research assistance.

## 1. Introduction

Statistics on the economically active population, employment, and unemployment are a key input in designing macroeconomic policies. They are an essential base for the design and evaluation of government programs geared towards employment creation, income enhancement, and poverty reduction. For policy responses to be meaningful, labor market statistics need to be reliably sourced, accurate, and timely. Typically, these statistics are generated from various sources. Data on employment can be obtained from establishment sample surveys, establishment or economic censuses, social security records, and public sector payrolls. Data on unemployment can be derived from administrative records on registered job seekers or recipients of unemployment benefits.

In India, historically, employment estimates have been generated using household and establishment surveys. Household surveys capture both the organized and unorganized sectors, particularly the self-employed. They largely satisfy the requirements of completeness. The quinquennial household surveys on employment and unemployment conducted by the National Sample Survey Office (NSSO), the last of which was conducted in 2011–12, have been the primary source of various labor market indicators since 1972–73. Establishment surveys, on the other hand, compile data from worksites and provide a more detailed picture of the industry structure of employment and characteristics of establishments. The key establishment surveys include the Economic Census (EC), the Annual Survey of Industries (ASI), and NSSO’s “Unincorporated Non-agricultural Enterprises (excluding construction).” While the EC provides the most comprehensive database of non-agricultural economic establishments in the country, the latter two are follow-up enterprise surveys of the EC. The ASI database is an annual one compiling information on the growth, composition, and structure of “registered” or formal sector firms, while NSSO’s Unincorporated Enterprise Survey is a quinquennial survey which provides data on unregistered firms.

Much of India’s employment data has been generated with a considerable time lag and often focuses on the organized sector. Recognizing the challenges arising from the paucity of real-time jobs data, the government set up a Task Force in 2017 to revamp the employment data architecture (NITI Aayog 2017). Significant among the recommendations of the 2017 Task Force was the introduction of the use of high-frequency administrative data (relating to payrolls, social security systems, and provident fund) and the discontinuation of NSSO’s quinquennial Employment and Unemployment

Surveys (EUS). NSSO's EUS have been replaced with the Periodic Labour Force Survey (PLFS), a household survey that will generate Key Indicators of the Labour Market (KILM), such as the labor force participation rate, worker population ratio, and unemployment rate (UR) at an annual frequency in rural areas and quarterly frequency in urban areas.

The revamp of employment data needs to be seen against the backdrop of the fact that employment in India is dominated by low-wage and low-productivity jobs in the unorganized sector. In addition to the dominance of the unorganized sector, evidence suggests that the organized sector has witnessed rapid informalization over the years through the contractualization and casualization of the workforce. What is more, many of the new jobs being created in the platform economy are also non-standard in nature and are outside the ambit of labor regulations. Thus, as old forms of informal employment persist, new forms are also emerging and there is an urgent need to improve statistics on the informal economy so that we have a nuanced understanding of the dynamics underlying this sector.

Additionally, India's jobs challenge is not just one of unemployment but also of underemployment. The UR, by itself, is an insufficient metric to anchor the policy discourse. In developing countries, due to the lack of sufficient unemployment benefit schemes, labor underutilization less often takes the explicit form of unemployment, showing up more among the employed (time-related underemployment) and persons outside the labor force (the potential labor force; International Labour Organization 2018). In order to have a comprehensive picture of the state of the labor market, particularly of the extent of labor underutilization, it is crucial to complement the analysis on labor market indicators with indicators on the quality of employment.

Given this backdrop, this paper will begin by examining the key characteristics of India's labor markets in Section 2. In Section 3, we examine the current state of data collection on employment and unemployment. Thereafter, in Section 4, we discuss the inadequacies of the existing data architecture in dealing with the complexities of the Indian setting. While we identify the gaps in data collection and analysis, we also make a series of recommendations for generating more reliable and relevant labor market data. Through these recommendations, we hope to move towards a more appropriate framework and better protocols to collect employment data—one that is cognizant of the formal and the informal sides of the labor market. Strengthening the statistical base of the world of work and achieving accuracy in capturing the different dimensions and nuances of the labor market are essential for shaping sound economic policy.

## 2. Characteristics of India's Labor Markets

Before delving into a discussion on labor market statistics, it is important to have a conceptual understanding of the nature of India's labor market conditions and how these have evolved over time. This has directional implications for both data collection and analysis. Many are inclined to think that all we need to know about the employment situation in an economy is how many are employed and how many are unemployed. And that we need to know these with high frequency. However, given the complexities of the Indian labor market, we need to know much more. Unemployment and employment rates are inadequate metrics to understand India's labor markets for multiple reasons. We will now examine why this is the case.

To begin with, open URs in India have typically been quite low. Table 1 reports the URs for India since 2004–05. It is worth noting that barring the recent PLFS, URs have fluctuated in the range of 2–3 percent, suggesting that India does not face a significant employment challenge. However, a disaggregated look at the UR suggests that this is not the case. An examination of the UR by education levels in Table 2 shows that for those with relatively higher education, the UR is quite high. For instance, for urban males and females who had received secondary education and above, the UR stood at 4 percent and 10 percent, respectively, in 2011–12, jumping to 9.2 percent and 19.8 percent, respectively, in 2017–18. This is largely a result of the fact that in an economy where the cost of remaining unemployed is very high, in the absence of social security, unemployment is a luxury enjoyed by the relatively well-off and educated individuals who can afford to remain unemployed (Ghose 2016). In contrast, the UR of those who were not literate

**TABLE 1. Percentage Unemployment Rates according to Usual Status (Principal Status (PS) + Subsidiary Status (SS)) and Current Weekly Status (CWS)**

Round (Year)	Rural				Urban			
	Male		Female		Male		Female	
	Usual Status		Usual Status		Usual Status		Usual Status	
	(PS + SS)	CWS						
PLFS (2017–18)	5.8	8.8	3.8	7.7	7.1	8.8	10.8	12.8
68 <sup>th</sup> (2011–12)	1.7	3.3	1.7	3.5	3.0	3.8	5.2	6.7
66 <sup>th</sup> (2009–10)	1.6	3.2	1.6	3.7	2.8	3.6	5.7	7.2
61 <sup>st</sup> (2004–05)	1.6	3.8	1.8	4.2	3.8	5.2	6.9	9.0

Source: Ministry of Statistics and Programme Implementation (2019, Statement 31).

**TABLE 2 . Percentage Unemployment Rates according to Usual Status (PS + SS) for Persons Aged 15 Years and above by Educational Attainment**

<i>General Education Level</i>	<i>Rural Male</i>			<i>Rural Female</i>		
	2004-05	2009-10	2011-12	2004-05	2009-10	2011-12
Not literate	0.3	0.3	0.5	0.2	0.0	0.2
Literate & up to primary	1.0	1.0	1.0	1.1	0.5	0.3
Middle	1.6	1.8	1.8	3.4	2.3	2.5
Secondary & above	4.4	3.5	3.6	15.2	11.8	9.7
All	1.6	1.6	1.7	1.8	1.6	1.6
	<i>Urban Male</i>			<i>Urban Female</i>		
Not literate	1.0	1.0	0.7	0.3	0.9	0.4
Literate & up to primary	2.1	1.6	1.9	2.9	0.5	1.3
Middle	4.2	2.6	2.2	8.0	3.7	3.0
Secondary & above	5.1	3.6	4.0	15.6	12.2	10.3
All	3.7	2.8	3.0	6.9	5.7	5.3

Source: Ministry of Statistics and Programme Implementation (2019, Statement 32).

fluctuated at around 1 percent. These are the poor who cannot afford to remain unemployed. They are compelled to resort to low-productivity and low-paying work in the unorganized sector. In 2017–18, 81.3 percent of all workers were estimated to be working in the unorganized sector.<sup>1</sup> While employment in the unorganized sector has declined over time, from 88.7 percent in 2004–05, it continues to be large. In fact, labor markets in India can be best characterized by their dualistic structure with the prevalence of an organized sector that co-exists with a large “unorganized sector.”

Importantly, self-employment and casual wage employment are the main forms of employment in the unorganized sector.<sup>2</sup> Data from the NSS household surveys and the PLFS indicate that these two forms of employment have continued to account for a disproportionately large share of total employment. As of 2017–18, over 75 percent of the total workers were engaged in these two forms of employment and were thus outside the ambit of standard employer–employee relationships. A mere 23.3 percent were in Regular Wage Salaried (RWS) employment. Ghose (2016) attributes the dominance of self-employment and casual wage employment to the fact that they facilitate work-sharing arrangements. In self-employment, the working members of the household share the work and income from the household enterprise, while in casual wage employment, workers share the amount of wage employment available. It is “this feasibility of work sharing that makes it possible for the unorganized sector to function as a reservoir of surplus labor, which exists in the form of underemployment of many workers rather than in the form of unemployment of some workers” (Ghose 2016, p. 21). In a situation where most workers are underemployed, escaping unemployment

1. Using information on enterprise type reported in household surveys, estimates on the unorganized sector are computed by including all enterprises other than government establishments and public enterprises, all enterprises in the private corporate sector and those private non-corporate enterprises that employ at least 10 regular employees. As defined in Section 4, informal or unorganized enterprises include the household sector as household enterprises or, equivalently, unincorporated enterprises owned by households.

2. In the NSSO, persons who operate their own farm or non-farm enterprises or were engaged independently in a profession or trade on own account or with one or a few partners were deemed to be self-employed in household enterprises. The essential feature of the self-employed is that they have autonomy (decide how, where, and when to produce) and economic independence (in respect of choice of market, scale of operation, and finance) for carrying out their operation. A person who is casually engaged in others’ farm or non-farm enterprises (both household and non-household) and, in return, receives wages according to the terms of the daily or periodic work contract is considered as a casual laborer. RWS employees are persons who work in others’ farm or non-farm enterprises (both household and non-household) and, in return, receive salary or wages on a regular basis (i.e., not on the basis of daily or periodic renewal of work contract). This category includes not only persons getting time wage but also persons receiving piece wage or salary and paid apprentices, both full-time and part-time.

means little. We need to have a better and deeper understanding about conditions of employment and whether the jobs they are engaged in are decent jobs. In other words, do these jobs provide for satisfactory working and living conditions and are people in employment necessarily better off than the unemployed.

Apart from its dualistic nature, another striking feature of India's labor markets has been the increasing informalization of the organized sector over time. Although the share of the organized sector has increased, albeit at a modest pace, the share of formal employment (defined by a worker having access to at least one social security benefit) has declined. In 2004–05, the share of regular formal employment in total employment in the organized sector stood at 51.3 percent. It declined to 47.4 percent in 2017–18.<sup>3</sup> Trends from the ASI also corroborate this phenomenon. Between 2000–01 and 2016–17, approximately half the increase in total employment, from 7.7 million to 13.7 million, was accounted for by the increasing use of contract workers. The share of contract workers in total employment increased sharply from 15.5 percent in 2000–01 to 27.9 percent in 2015–16, while the share of directly hired workers fell from 61.2 percent to 50.4 percent during the same period. It is worth mentioning that even amongst RWS workers, conditions of employment leave much to be desired. As of 2017–18, 71.1 percent of RWS workers had no written job contract and 49.6 percent were not eligible for social security benefits. The contractualization, casualization, and informalization of the workforce point to the need for capturing changes in the quality of employment data.

In such a scenario, simply having a count of how many people are unemployed and employed, and what the employment status of the latter is, is clearly inadequate. Neither the current state of employment conditions nor the changes over time can be discerned simply from observed changes in employment and unemployment rates. To know if employment conditions are improving, we need to know if workers are moving from low-productivity to high-productivity work and if the level of underemployment is declining.

### 3. The Current Data Architecture

At present, the sources of statistics on India's employment indicators can be grouped into three broad categories: household sample surveys;

3. Regular formal employment is defined as RWS employment that offers at least one social security benefit.

establishment censuses and establishment sample surveys; and various types of administrative records. These sources differ in the type and detail of information they provide; in coverage and periodicity; in concepts, definitions, and measurement units; and in cost of operation, quality, and timeliness of the results. We will briefly describe each of these.

### 3.1. Household Surveys

To get a holistic picture of India's dualistic labor markets, household surveys are the most widely used to generate employment estimates. Given their ability to capture both the organized and unorganized sectors, particularly household enterprises, they provide the most comprehensive data on the employment situation in the country. The main objective of EUS, which have been conducted by NSSO at periodic intervals since 1972–73, is to get estimates of level parameters of various labor force characteristics at the national and state/Union Territory (UT) levels. The participation of people is not only dynamic but also multidimensional—it varies with region, age, education, gender, industry, and occupational category. These aspects of the labor force are captured in detail in NSSO's EUS and estimates are generated for the labor force participation rate, worker population ratio, UR, extent of underemployment, and wages of employees. The key indicators of the labor market and how they are derived are depicted in Table 3. The indicators of the structural aspects of the workforce, such as status in employment, industrial distribution, and occupational distribution are also derived from these surveys.

The detailed activity statuses under each of the three broad activity statuses (viz., “employed,” “unemployed,” and “not in the labor force”) and the corresponding codes used in the survey are reported in Table 4.

Following the recommendations of the “Expert Committee on Unemployment Estimates” (popularly known as the Dantwala Committee 1970), which noted that no one-dimensional measure was meaningful in the

**TABLE 3. Key Indicators of the Labor Market Used by NSSO**

---

1. <i>Labor force participation rate (LFPR)</i> : [(No. of employed persons + No. of unemployed persons) / Total population] × 1,000
2. <i>Worker Population Ratio (WPR)</i> : [No. of employed persons / Total population] × 1,000
3. <i>Proportion of Unemployed (PU)</i> : [No. of unemployed persons / Total population] × 1,000
4. <i>Unemployment Rate (UR)</i> : [No. of unemployed persons / (No. of employed persons + No. of unemployed persons)] × 1,000

---

Source: National Sample Survey Office (2015).

**TABLE 4. NSSO Activity Status Codes for Employed, Unemployed, and Not in Labor Force Categories**

<i>Working (Employed)</i>		
<i>Self-employed</i>	<i>Regular Wage/ Salaried Employee</i>	<i>Casual Labor</i>
11: Worked in household enterprises (self-employed) as own-account workers; 12: Worked in household enterprises (self-employed) as employers; 21: Worked in household enterprises (self-employed) as helpers	31: Worked as regular wage/salaried employees	41: Worked as casual laborers in public works other than MGNREGA public works; 42: Worked as casual laborers in MGNREGA public works; 51: Worked as casual laborers in other types of works; 61: Did not work owing to sickness though there was work in household enterprises; 62: Did not work owing to other reasons though there was work in household enterprises; 71: Did not work owing to sickness but had regular salaried/wage employment; 72: Did not work owing to other reasons but had regular salaried/wage employment.
<i>Not Working but Seeking/Available for Work (Unemployed)</i>		
81: Sought work or did not seek but were available for work (for Usual Status approach); 81: Sought work (for Current Weekly Status approach); 82: Did not seek but were available for work (for Current Weekly Status approach).		
<i>Neither Working nor Available for Work (Not in the Labor Force)</i>		
91: Attended educational institutions; 92: Attended to domestic duties only; 93: Attended to domestic duties and were also engaged in free collection of goods (vegetables, roots, firewood, cattle feed, etc.), sewing, tailoring, weaving, etc., for household use); 94: Rentiers, pensioners, remittance recipients, etc.; 95: Not able to work owing to disability; 97: Others (including beggars, prostitutes, etc.); 98: Did not work owing to sickness (for casual workers only); 99: Children of age 0–4 years.		

Source: National Sample Survey Office (2015).

Note: Activity Status Codes 42, 61, 62, 71, 72, 82, and 98 are used only in the current status (in CWS and CDS).

Indian context, the NSSO has been producing three types of estimates of the employed and the unemployed. These are the Usual Status (US), Current Weekly Status (CWS), and Current Daily Status (CDS). The methodology for assigning the three statuses is different, and therefore, their interpretation also differs.

The estimate of the employed, according to the Usual Principal Status, gives the number of persons who worked for a relatively longer part of the

reference period of 365 days preceding the date of the survey. The workforce, considering both the Usual Principal Status and the Subsidiary Status (together referred to as Usual Status), includes the persons who: (a) either worked for a relatively longer part of the 365 days preceding the date of the survey and (b) were among the remaining population that had worked for a shorter time throughout the reference year of 365 days preceding the date of the survey or for a minor period, which is not less than 30 days. The estimate of the workforce, according to the CWS, includes those who worked for at least one hour on any day during the seven days preceding the date of the survey. The workforce measured in terms of the CDS gives the average picture of the number of person-days worked during the survey period. For each person, seven person-days are assigned for the week preceding the date of the survey. Each day of the reference week is seen as comprising either two “half days” or a “full day” for assigning the activity status. For recording the time disposition for activities pursued by a person in a day, an intensity of 1.0 is given against an activity that is done for “full day,” and an intensity of 0.5 against the activity which is done for “half day.” A person is considered as “working” (employed) for the full day if he/she had worked for four hours or more during the day.

Given the exhaustive and comprehensive nature of its coverage, data from the NSSO’s EUS is collected only once every 5–6 years. Another drawback of these surveys is their inability to provide reliable estimates of employment below the state level, that is, the district level. In fact, there are precision issues even for smaller states and UTs. There is also a high margin of errors associated with the UR by gender at the state level in some cases. It is important to mention that NSSO’s consumption surveys conducted on an annual basis (i.e., the thin rounds which are undertaken during the intervening period between the quinquennial rounds) had an employment schedule and were conducted until 2003–04. Thereafter, the NSSO conducted several thin rounds of the EUS between the quinquennial rounds of 2004–05 and 2011–12, before the survey was finally discontinued.<sup>4</sup> Realizing the need for regular and frequent labor statistics, the Labour Bureau also began conducting the EUS annually since 2009–10. It has conducted six surveys

4. NSSO’s annual thin sample is about 40 percent of the sample of households in the quinquennial survey. Although the annual and quinquennial surveys follow identical concepts, schedules of quinquennial surveys provide for detailed probing, which is not available in the annual rounds. There have been some differences in the estimates thrown by the annual and quinquennial data. NSSO does not give any indication of the reliability of its estimates, some of which may be very unreliable due to poor representation in the sample.

since then.<sup>5</sup> However, the results of the last survey for the year 2016–17 have not been released.

Following the recommendations of the 2017 Task Force, both the above-mentioned household surveys have been discontinued and replaced by the PLFS. The report of the first PLFS for the period 2017–18 was released in May 2019. Conducted by the NSSO, the PLFS produces annual estimates of KILM based on the US and CWS approaches. Additionally, in urban areas, a rotational panel sampling design is being used to generate quarterly urban estimates according to the CWS approach. Each selected household in urban areas is visited four times—in the beginning with the first visit schedule and thrice periodically later with the revisit schedule. There is no revisit in the rural samples.

In addition to the introduction of the quarterly module, the PLFS has some other noteworthy features. First, information in the schedule of enquiry has been collected from the sample households using the Computer-assisted Personal Interviewing (CAPI) method. Second, information is collected on the number of hours worked on each day of the reference week. For the persons engaged in economic activity on the days of the reference week, the number of hours worked is recorded and the current weekly activity status for those engaged in economic activity is determined on the basis of time disposition during each day of the reference week (recorded in terms of the number of hours worked). Besides, information is collected on the total number of hours actually worked considering all the work performed during the day. Third, for the persons who are engaged in economic activity on different days of the week, if they are available for additional work on a particular day, information on the number of hours available for work is recorded. Fourth, information on earnings from employment is collected. For RWS employees in CWS, information on earnings for the preceding calendar month is collected. For the self-employed persons in CWS, information on gross earnings from self-employment activity during the last 30 days is collected. For the persons who worked as casual laborers on different days of the reference week, information on wage earnings is collected separately for each of those specific days.

In a departure from NSSO's previous EUS, the recently released report of the PLFS does not provide any estimates using the CDS. The PLFS schedule includes a set of questions on current weekly activity particulars

5. The reference period for the Labour Bureau's surveys is as follows: 2009–10 (April 2009–March 2010); 2011–12 (July 2010–June 2011); 2012–13 (October 2012–March 2013); 2013–14 (January 2014–July 2014); and 2015–16 (April 2015–December 2015).

of the household members, which is similar to the questions posed on time disposition during the week in NSSO's EUS questionnaire. While researchers could possibly compute the CDS measures from the unit data, the reason for dropping the CDS measure from the discussion is not clear.<sup>6</sup>

Before concluding the discussion on data compiled from households, it is important to mention that the decennial population Census also collects information on the economic activity of the population. The economic tables of the Census (referred to as the B-Series Tables) provide information on the number of main workers (those who worked for the major part of the reference period, i.e., six months or more); marginal workers (those who did not work for the major part of the reference period, i.e., less than six months), and non-workers (those who did not work at all during the reference period). The main and marginal workers are also classified into four categories based on the economic activity performed by them during the last one year. These are cultivators, agricultural laborers, workers in the household industry, and other workers. Tabulations are also available by religious communities, social groups, and education level. Main workers are also classified by industrial category, education level, age, sex, and religious community. These tabulations are made available at the national, state, district, and city levels. That the economic tables of the Census provide reliable estimates of the number of workers at the district level is noteworthy. Even NSSO's EUS does not provide reliable estimates at the district level as the sample size is not large enough. Despite this advantage of the Census, the above-mentioned tabulations have not attracted much attention in the academic literature. This is probably a consequence of the fact that there is a considerable delay in the release of the series, which makes them dated. For instance, the economic tables for the Census of 2011 were released only in 2017.

### *3.2. Census and Survey of Establishments*

In addition to household surveys, enterprise or establishment surveys, which compile data from the workplace, are an important source of employment data. Data collected from worksites provides a more detailed picture of the industry structure of employment and characteristics of enterprises. In household surveys, where the respondent is the household head or member (who may not be the worker in question), obtaining correct information of the characteristics of the enterprise in which the worker works is challenging (Papola 2014).

6. It is also not clear if the methodology for computing the CDS will be the same as before and if the time intensity approach of using 1 or 0.5 will continue in the PLFS now that the actual number of hours being worked on a particular activity are being recorded.

Although agriculture is a sector that employs a significant part of the rural workforce, enterprise surveys typically cover activities other than agriculture. As mentioned earlier, there are two key enterprise surveys in India. The first is the ASI by the Ministry of Statistics and Programme Implementation (MoSPI). It is the main source of industrial statistics in India and provides detailed information annually on the organized manufacturing sector (comprising activities related to manufacturing processes, repair services, gas and water supply, and cold storage). It was launched in 1960 with 1959 as the reference year and has continued for all years since then, except for 1972. The survey gathers information only on “registered” or formal sector firms that are covered under Sections 2m(i) and 2m(ii) of the 1948 Factories Act, that is, those firms that use electricity and hire 10 or more workers, and those that do not use electricity but nevertheless employ 20 or more workers.

The definition of a worker in the ASI includes all persons employed directly or through an agency. It includes those engaged not just in any manufacturing process but also in cleaning any part of the machinery or premises used for the manufacturing process or in any other kind of work connected with the manufacturing process or the subject of the manufacturing process. The total number of persons engaged is defined as production workers (sum of the workers hired directly and contract workers), supervisory and managerial staff, and all working proprietors and their family members who are actively engaged in the work of the factory even without any pay, and unpaid members of the co-operative societies who worked in or for the factory in any direct and productive capacity. Importantly, the number of workers is an average number computed by dividing the total number of man-days worked in the factory by the number of days the factory had worked during the reference year. The total number of man-days worked during a month is, in turn, obtained by summing up the number of workers in each shift over all the shifts worked on all working days during a month. Thus, the number of workers obtained from the ASI would be the actual labor input in the manufacturing process rather than the actual number of individuals who worked in the factory. As such, comparing this number with estimates of the organized sector workforce from administrative sources or the EUS would not be conceptually valid. We will elaborate on the problem in comparing estimates of workers from enterprises and household surveys later.

The second main enterprise survey is the NSSO’s Enterprise Survey of Unincorporated Enterprises. The surveys are conducted quinquennially and have typically covered the manufacturing sector. Since 2010–11, they

have expanded their coverage to include trade and other service sectors (excluding construction). The NSSO classifies unregistered firms into three categories: (a) own-account manufacturing enterprises, that is, those that operate without any hired worker employed on a fairly regular basis, (b) non-directory manufacturing establishments, that is, those that employ fewer than six workers (household and hired workers taken together), and (c) directory manufacturing establishments, that is, those that employ a total of six or more household members and hired workers. The significance of this survey stems from the fact that it takes into account the self-employed and employment in establishments with less than 10 workers, which the ASI does not take into account.

In this survey, a worker is defined as one who “participates either full time or part time in the activity of the enterprise in any capacity—primary or supervisory—and may or may not receive wages/salaries in return.” The average number of persons usually working on a working day during the reference month is recorded. A worker refers to a position rather than a person. It includes working owners, hired workers (full-time and part-time), apprentices (paid/unpaid), and other workers/helpers working without regular salary or wages. In the case of proprietary or partnership enterprises, the owner(s) personally working in the enterprise on a fairly regular basis are treated as working owners.

In addition to the above-mentioned enterprise surveys, the Central Statistics Office under MoSPI conducts the EC. This is a count of all establishments/units engaged in the production of goods and services and is the most comprehensive database of non-agricultural economic establishments in the country. The most recent EC was conducted in 2013–14. Prior to this, it was conducted in 2005–06. While the main purpose of the EC is to provide a frame for other data collection exercises such as the NSS’s Unincorporated Non-agricultural Enterprise Surveys, it also provides basic information on the number of establishments/units, their employment location, type of activity, and nature of operation.

*PROBLEMS IN COMPARING ESTIMATES OF WORKERS FROM ENTERPRISE AND HOUSEHOLD SURVEYS*

Before proceeding, it is important to discuss the issue of comparability of household labor force surveys and enterprise surveys such as the ASI. In the Indian context, given the irregularity of employment patterns, the informal and transitory nature of enterprises, and differences in the concepts and procedures used in the two surveys, we cannot make a direct comparison between them. This can be explained as follows.

First, establishment surveys count the number of positions (jobs) as opposed to the number of employed persons. As explained above, in establishment surveys, the average number of workers is calculated by counting the total number of man-days worked by different workers and dividing this by the total number of working days.<sup>7</sup> As against this, in the EUS, a person is termed as employed (worker) if he or she is engaged in economic activities on the criterion of major time during a pre-specified reference period.

Second, we need to make a distinction between the reference periods of enterprise and household surveys. Many enterprises operate for less than six months and some of them operate for even less than three months. In enterprise surveys, all the persons employed in enterprises are counted as workers, making no allowance for the duration. On the other hand, in the EUS, the reference period for the US criteria is taken as a year. The US criteria for these entrepreneurs will be decided on the basis of the activity pursued by them during the remaining part of the year. They will be considered as employed only in the subsidiary status if they have not followed any other economic activities during the major period.

Third, some workers tend to get excluded from enterprise surveys. For instance, workers who work for households rather than for enterprises (such as domestic servants and drivers) get left out of enterprise surveys.<sup>8</sup> Casual laborers who cannot be assigned to an enterprise also do not get captured in the enterprise survey as these surveys cover those workers (be it part-time or full-time) who participate in the activities of the enterprise on a “fairly regular basis.”

Fourth, the treatment of part-time workers poses a problem. According to the NSSO’s enterprise survey instructions, full-time workers are those who work for more than half of the normal working hours on a fairly regular basis, while part-time workers are those who work for less than half the normal working hours of the enterprise on a regular basis. In the enterprise survey, two part-time workers will be treated as two and not one full-time worker leading to an overestimation of the number of workers.

### *3.3. Administrative Data*

Following the recommendations of the 2017 Task Force, the government has started reporting administrative data relating to payrolls, social security systems, and provident fund for compiling information about the

7. This is then rounded off to a whole number.

8. In enterprise surveys, it is operationally difficult to list all employer households as enterprises.

labor market. The first such series was released in January 2018 pertaining to the time period, April 2017–November 2017. Since then, the series has been released on a monthly basis providing age-wise payroll data for the Employees’ Provident Fund Organisation (EPFO), Employees’ State Insurance Corporation (ESIC), and National Pension Scheme (NPS). Payroll databases are widely used in advanced economies to gauge the employment situation and provide a count of formal jobs. The introduction of these databases in India is indeed a significant initiative. However, the interpretation and analysis of India’s payroll data is not so simple. As outlined in Kapoor (2018), it is an exercise fraught with challenges for multiple reasons.

First, there is significant overlap and duplication across the above-mentioned schemes. We need a common identifier across the multiple datasets to avoid double counting. Attempts are being made to circumvent this problem using the unique *Aadhaar* identity number. While the EPFO and Pension Fund Regulatory and Development Authority datasets have been *Aadhaar* seeded, the ESIC data has not. Second, new entries on these databases do not necessarily reflect new jobs. Given that the EPFO Act applies to all factories in classes of industry specified in Schedule 1 of the Act where 20 or more persons are employed, the addition of even one more worker will result in all the workers of this establishment getting added on the database and being counted as new jobs. Despite the 2017 Task Force pointing to this caveat, these numbers are widely being interpreted as indicative of new formal jobs.

Third, each time the data has been released, the previously released estimates have been revised. This constant volatility begs the question as to whether this data is in fact reliable and accurate in “real time.” Conceptually, too, the EPFO database has seen a significant revision. In August 2018, the EPFO started reporting the number of subscribers who ceased their subscriptions and later “rejoined” and “re-subscribed” to the database. This category has now been included in the net additions to databases implying that those who left a formal job and rejoined a formal job are counted as new formal jobs. These revisions not only highlight the fragility of this database, but they also give the sense that this database is still in experimental mode.

Given the above limitations, the excessive focus on the payroll data to understand employment trends serves little purpose. The larger issue, of course, is that it cannot provide a holistic picture of the employment scenario as it focuses simply on the formal sector. Nevertheless, it is important to undertake administrative reforms to strengthen and develop the payroll database such that it can emerge as an important data source in the future.

### *3.4. Private Data Sources*

To fill in the gaps in India's official statistical employment database, the Centre for Monitoring Indian Economy (CMIE), a private agency, has also begun tracking unemployment since January 2016. The CMIE has been compiling an unemployment database that contains detailed information on the employment and unemployment status of the members residing in all households in their Consumer Pyramids sample. The survey, referred to as the Consumer Pyramids Household Survey (CPHS), has a sample size of more than 161,000 households covering about 522,000 people. This makes its sample size larger than that of NSSO's EUS. Employment estimates are based on consecutive waves of the CPHS by tracking a panel of households. Each wave is executed over a period of four months and there are three waves in a year. Complete demographic data of all households in the Consumer Pyramids sample is collected thrice a year. The employment status of the members of the households is recorded as one of the following:

- Employed
- Unemployed, willing to work and actively looking for a job
- Unemployed, willing to work and not actively looking for a job
- Unemployed, not willing to work and not actively looking for a job

Given the high-frequency statistics it generates, the CMIE database has emerged as an important source of employment data. Nevertheless, it cannot be seen as a substitute for the comprehensive household surveys conducted by the NSSO, which provide an in-depth understanding of labor's engagement in economic activities. Conceptually, the surveys are different too. While the NSSO provides employment estimates using three different approaches based on different reference periods, the CMIE survey seeks a response on the employment/unemployment status from every member of the sample household, who is aged 15 years or more on the day of the survey. If the status on the day of the survey is unclear (such as for a daily wage worker), then the status as of the previous day is taken instead. Such an approach helps to ensure that there is no ambiguity regarding status and no problem regarding recall. An important feature of the CMIE survey is the use of technology which enables them to generate estimates at high frequency. Data are captured on a mobile phone on a specially developed software application and validated and uploaded for use by the end of day, every day. This has enabled CMIE to generate estimates on a daily, weekly, and monthly basis. While this is a welcome development, data from private agencies cannot be seen as a substitute for official government data.

## 4. Issues and Recommendations

The preceding section explains how India's employment data architecture has changed since the original systems were introduced. Despite these changes, significant gaps exist in data collection and analysis. In this section, we examine these gaps and understand how they can be addressed to get a more meaningful picture of India's dualistic labor markets. We provide feasible solutions where possible. Significantly, these recommendations (summarized in Table 5) do not entail conducting new surveys but making more effective use of the data being currently generated.

### 4.1. *The Relevance of Measuring Indicators using Three Approaches*

The move from NSSO's quinquennial EUS to the annual PLFS requires us to revisit the issue of the approach to adopt in measuring employment. India's household surveys have produced employment measures using three approaches—Usual Status, Current Weekly Status, Current Daily Status. The policy debate thus far has been anchored around US measures.

**TABLE 5. Summary of Key Suggestions**

1. Continue with the NSS quinquennial EUS round while concomitantly holding the annual PLFS.
2. In addition to the UR, create three new measures of labor underutilization in line with the recommendations of the 19<sup>th</sup> ICLS: combined rate of underemployment and unemployment; combined rate of unemployment and potential labor force, and composite measure of labor underutilization. These three measures should be included amongst the KILM reported in all the survey results.
3. Re-introduce the CDS measure in the PLFS to understand the extent of underemployment and daily changes in activities.
4. Incorporate additional questions on worker classification to understand and categorize different work arrangements such as contract workers, fixed-term employees, gig-economy workers, and homeworkers.
5. Create a headline rate of informal employment which combines the enterprise-based definition of informal sector with the job-based definition of informal employment to provide a more holistic picture of the extent of informality in the economy. This measure should also be included in the list of KILM provided in the PLFS reports.
6. Supplement the quarterly module in urban areas with a quarterly module in rural areas and expand the sample size.
7. Implement a National Business Register and ensure that states complete their State Business Register.
8. Present results in a more user-friendly and accessible manner.
9. Leverage technology to reduce the time lag in the collection and dissemination of data.

Source: Author's summary.

As it pertains to the activity status of a person during the reference period of 365 days preceding the date of the survey, its interpretation is fairly uncomplicated.

However, as the PLFS moves towards providing annual estimates of KILM at the national and rural levels, and quarterly estimates in urban areas, it may no longer be appropriate to base policy discussions on US measures. Given the predominance of self-employment and large-scale employment in the agricultural sector, the US measures are not expected to fluctuate over short periods. Being inherently smoother, they are unlikely to capture changes in activity patterns caused by seasonal fluctuations and witness much variation over short periods. Given this shortcoming of the US approach, the PLFS estimates employment and unemployment using the CWS concept as well. For the urban quarterly estimates, it reports only the CWS measures. The increased importance of CWS in the PLFS is indeed a positive development, as this is the most widely used approach across various national labor force surveys.

However, it is important to mention that there is a challenge in interpreting the CWS. In all NSS surveys, the reference period is fixed in relation to the date of the survey. For the US, an entire year is referenced, though the year may be different for different persons in the sample. To obtain the rate of unemployed or the workforce participation, we consider the total number of persons getting classified as unemployed or employed in the numerator, while the denominator is the estimated total number of persons in the labor force or the relevant population cohort. There is no problem in understanding such a measure as an average rate for the persons in the population. However, the CWS cannot be interpreted in a similar manner (Mohanani 2019a). In the CWS, for each person, we consider a particular week out of the 52 weeks in the year. Different weeks are considered for different persons. Based on the criteria outlined in the concept of CWS, a week is classified as an “employed” week, “unemployed week,” or “out of labor force week.” Thus, we have a large number of weeks classified according to these three criteria, though they pertain to different persons.

The question that then arises is whether the status of a single week of a person can be added to that of another person for another week and divided by the number of persons to obtain any meaningful average. Given the diversity of employment situations in India and the fact that the status of persons changes during the year frequently for a very large segment of the population, it would not be correct to equate the status obtained on the basis of a single week in the year/quarter to the status of the person during the survey period. Therefore, the rate cannot be called a rate for persons. Instead,

it can be referred to as the unemployed (employed) person-weeks (Mohanam 2009). A similar problem arises in the interpretation of the CDS as it gives ratios in terms of person-days. Thus, these two measures are conceptually not suited to building up estimates of either employed or unemployed persons. However, this does not mean that we should not use these measures. On the contrary, they are particularly important in the Indian context.

In fact, the CDS approach, which has been dropped in the PLFS, provides the most comprehensive and detailed picture of the employment situation by reporting a day-to-day account of the available labor time during the week. Understanding these day-to-day changes of the week is particularly useful for workers engaged in informal employment as their activity pattern is not expected to be the same every day. Typically, the appropriate choice is between the week or the day as the reference period depends on national conditions, particularly on the extent of weekly and daily fluctuations of activity status. Where the dominant form of employment is regular full-time employment, the “week” should be the preferred reference period. In such employment conditions, a reference week or a reference day (of the same week) will generally give similar average results, but the advantage of choosing the week is that it will lead to results with lower variances. Where a week of employment does not generally represent a whole week of full-time employment, that is, where casual and intermittent work and part-time and temporary jobs and other types of short-term employment are widespread, then the one-day reference period may be preferable, since it would give a sharper snapshot of the employment and unemployment situation of the country than a weekly measure. The International Labour Office (Husmanns, Mehran, and Verma 1990) notes that in economies such as India, where casual and intermittent work are widespread, the daily measure would give a clearer snapshot of the employment scenario. Further, the Report of the Special Group on Targeting Ten Million Employment Opportunities per Year (Planning Commission of India 2002) has also noted that CDS is a better measure than the US, as the latter includes the underemployed workforce and gives a misleading picture to policymakers about the extent of labor underutilization.

Having discussed the pros and cons of all three measures, the question of which actually works best arises. Given the preponderance of informality, wherein a major section of the workforce works in the household sector, while a small proportion are engaged in highly formal employment, NSSO surveys have devised multiple reference periods and multiple concepts of employment and unemployment. These have served their purpose well and it is pertinent to continue with all three approaches going forward. It is also

important to bring back the CDS measure in the discourse and clarify that CDS and CWS measures need to be interpreted in terms of person-days and person-weeks, respectively.

#### *4.2. Measuring Labor Underutilization*

Labor underutilization refers to mismatches between labor supply and demand, leading to an unmet need for employment among the population. While unemployment is at the core of the concept of labor underutilization, the UR by itself provides only a partial view of the labor underutilization in an economy such as India where the latter often takes the form of underemployment. According to the 1996 resolution of the 11<sup>th</sup> International Conference of Labour Statisticians (ICLS), underemployment “exists when a person’s employment is inadequate, in relation to specified norms or alternative employment, account being taken of his occupational skill (training and working experience).”

There are two principal forms of underemployment: visible and invisible. The former reflects an insufficiency in the volume of employment and according to the 13<sup>th</sup> ICLS (1982), it is defined as comprising “persons in paid or self-employment, whether at work or not at work, involuntarily working less than the normal duration of work determined for the activity, who were seeking or available for additional work during the reference period.”<sup>9</sup> On the other hand, invisible underemployment is characterized by low income, underutilization of skills, and low productivity. Compared to visible underemployment, which is a statistical concept directly measurable in the labor force, invisible underemployment “is primarily an analytical concept reflecting a misallocation of labor resources or a fundamental imbalance as between labor and other factors of production” (ICLS 1982). The NSSO has also been attempting to compute measures of visible and invisible underemployment in its quinquennial household surveys. We will briefly outline how this has been done in the NSS 68<sup>th</sup> Round (2011–12), as follows.

1. *Visible Underemployment:* Some of the persons categorized as usually employed may not have work throughout the year due to seasonality in work or other reasons. Their labor time is not fully utilized and they are, therefore, underemployed. In the same way, persons categorized

9. For the purpose of classifying persons as visibly underemployed, the normal duration of work for an activity should be determined in the light of national circumstances as reflected in national legislation to the extent it is applicable, and usual practices in other cases, or in terms of a uniform conventional norm.

as employed in current weekly status may not have work for all the days of the reference week. Underemployment of such employed persons is termed as visible underemployment if they report themselves to be not employed with respect to a shorter reference period. Three measures of visible underemployment are derived by NSSO by classifying: (a) usually employed persons by their current weekly status, (b) usually employed persons by their current daily status, and (c) persons employed in current weekly status by their current daily status. This classification generates the following rates of visible underemployment:

- i. Proportion of usually employed persons who did not have work during the reference week
- ii. Proportion of person-days of usually employed persons that were not used for work
- iii. Proportion of person-days of persons employed in current weekly status that were not used for work

A cross-examination of the activity status of the population with respect to each of the three approaches can throw light on the extent of visible underemployment and provide a nuanced understanding of labor market underutilization. In fact, the need to capture visible underemployment reinforces the need to incorporate the CDS measures in the PLFS reports.

2. *Invisible Underemployment*: Some persons categorized as employed according to the framework adopted in NSSO's EUS may not have enough work or may feel that their labor potential is not fully utilized or that the work they pursue may not fully fulfill their requirements. Therefore, they may want additional and/or alternative work. Such underemployment is termed as invisible underemployment and is not directly measurable. To estimate invisible underemployment, NSSO's EUS questionnaire posed a set of probing questions to the usually employed (Table 6 reports these questions which appeared in Block 6 of the NSS Schedule of Enquiry). The questions pertained to whether the usually employed were available for additional/alternative work and the reason for seeking such additional/alternative work (Columns 10–13 of Block 6 of the NSS Schedule). More specifically, those seeking additional work were asked whether they did so to supplement their incomes and/or they did not have enough work or there were some other reasons (Column 11). Similarly, those seeking alternative work were asked if they did so because their present work was not remunerative enough, provided no job satisfaction, lacked job security,

**T A B L E 6 . Block 6 of the NSSO Schedule of Enquiry for the Employment Unemployment Survey (2011–12): Follow-up Questions on Availability for Work, Existence of Union/Association and Nature of Employment for Persons Working in the Usual Principal or Subsidiary Status (i.e., those with codes 11–51 in Column 3 of Block 5.1 or Block 5.2)**

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
srl. no. as in col.1, Block 5.1	age (yrs.) as in col.2, Block 5.1	Usual Activity Status Code	Principal (as in col.3, Block 5.1)	Whether Engaged Mostly in Full Time or Part Time Work during Last 365 Days (Full time -1, Part time -2)	Whether Worked More or Less Regularly during Last 365 Days (Yes-1, No-2)	Approximate No. of Months without Work (Months)	If Entry > = 1 in Col. 7, Whether Sought / Available for Work during Those Months (Code)	For Codes 1 & 2 in Col. 8, Whether Made Any Efforts to Get Work (Code)	Whether Sought/ Available for Additional Work during the Days He/She Had Work (Code)	For Codes 1 & 2 in Col. 10, Reason (Code)	Whether Sought/ Available for Alternative Work during the Days He/She Had Work (Code)	For Codes 1 & 2 in Col. 12, Reason (Code)	Is There Any Union/ Association in Your Activity? (Yes -1, No -2, Not Known -9)	For Code 1 in Col. 14, Whether a Member of Union/ Association (Yes-1, No-2)	Nature of Employment (Permanent-1, Temporary-2)

Source: National Sample Survey Office (2015).

Notes: Codes for Block 6

col. (8): whether sought/available for work during those months: yes: on most days -1, on some days -2; no -3.

col. (9): whether made any efforts to get work: yes: registered only in government employment exchanges-1, registered only in private placement agencies-2, registered in both government employment exchanges and private placement agencies-3, other efforts -4; no effort-5.

col. (10): whether sought/available for additional work during the days he/she had work: yes: on most days -1, on some days -2; no -3.

col. (11): reason for seeking/available for additional work: to supplement income -1, not enough work -2, both -3, others -9.

col. (12): whether sought/available for alternative work during the days he/she had work: yes: on most days -1, on some days -2; no -3.

col. (13): reason for seeking/available for alternative work: present work not remunerative enough -1, no job satisfaction -2, lack of job security -3, workplace too far -4, wants wage/ salary job -5, others -9.

the workplace was too far, they desired a wage/salary job instead, or there were other reasons<sup>10</sup> (Column 13). The proportion of the usually employed who reported their availability for additional/alternative work gives an overall share of the employed who did not have enough work, or at least felt the need for additional/alternative work, and, in that sense, were underemployed. Two other important follow-up questions posed to the usually employed, which helped estimate invisible underemployment, were: (a) whether they worked more or less regularly during last 365 days (Column 6), and (b) whether they were engaged mostly in full-time or part-time work during the reference period of 365 days (Column 5). The EUS also asked the usually employed the approximate number of months they spent without work<sup>11</sup> (Column 7). Mohanan (2009) has undertaken a detailed analysis using these questions to understand the nature of underemployment amongst the usually employed for NSSO's EUS 61<sup>st</sup> Round (2004–05).

Although the above-mentioned questions offer important insights into the nature of usual status employment and reasons for the usually employed seeking additional/alternative work, the PLFS has dropped them from the questionnaire. Instead, it only asks workers if they were available for additional work and how many hours they were available for such work during the day. The PLFS Report provides details on the percentage of workers in the CWS who were available for additional work. It provides this distribution by the employment status of the worker (i.e., self-employed, RWS employees, and casual laborers) on a quarterly basis. While this information can potentially give us a sense of time-related underemployment, simply providing the number of hours for which they sought additional work without understanding the reason for doing so offers little insight into the nature of the underemployment.

In order to get an estimate of visible and invisible underemployment, we propose re-introducing the detailed questions in Table 5 to the PLFS questionnaire. While we understand that incorporating elaborate follow-up questions

10. For instance, in the NSS 68<sup>th</sup> Round (2011–12), of those who sought additional work, 59.6 percent in rural areas and 62.2 percent in urban areas said they did so to supplement their income. Approximately 15 percent in rural areas and 11.5 percent in urban areas said they sought additional work as they did not have enough work at present. The report also finds that of those who sought alternative work, 57.8 percent in rural areas and 59.5 percent in urban areas said their present work was not remunerative enough. Close to 15 percent sought alternative work as they had no job satisfaction in their current work.

11. These months were to be counted after deep probing and identifying all months without work even if they were isolated months and the days were rounded off to the nearest month.

in the PLFS, which is an annual survey, is difficult and cumbersome, the exclusion of these questions hinders our understanding of the nature of usual status employment in India. The continuation of NSSO's quinquennial EUS, which has an exhaustive and comprehensive questionnaire, is imperative to gain more detailed insights into the nature of employment, in particular, underemployment. It is worth noting that the PLFS was never meant to replace the thick quinquennial rounds, which had a more extensive coverage of issues. The latter were meant to be held every 4–5 years, while the PLFS was conducted concomitantly at an annual frequency in rural areas and at a quarterly frequency in urban areas.

We propose exploiting the information on invisible and visible underemployment, compiled by NSSO's EUS (through questions asked in Block 6) to create new measures of labor underutilization in line with the 19<sup>th</sup> ICLS's (Work Statistics Committee 2013) resolution on statistics of work, employment, and labor underutilization. Recognizing the limitations of the UR as a measure of labor underutilization, the resolution laid out the following four measures of labor underutilization:

1. LU1: Unemployment Rate =  $(\text{Unemployment} \times 100) / \text{Labor Force}$
2. LU2: Combined Rate of Time-related Underemployment and Unemployment =  $(\text{Time-related Underemployment} + \text{Unemployment}) \times 100 / \text{Labor Force}$

where time-related underemployment is defined as persons in employment whose working time is insufficient in relation to alternative employment situations in which they are willing and available to engage.

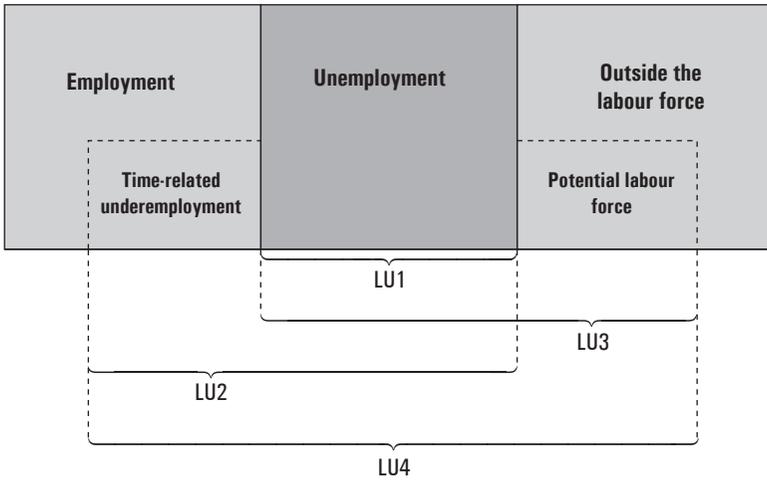
3. LU3: Combined Rate of Unemployment and Potential Labor Force =  $(\text{Unemployment} + \text{Potential Labor Force}) \times 100 / (\text{Labor Force} + \text{Potential Labor Force})$

where potential labor force is defined as persons not in employment who express an interest in it but for whom existing conditions limit their active job search and/or their availability.

4. LU4: Composite Measure of Labor Underutilization =  $(\text{Time-related Underemployment} + \text{Unemployment} + \text{Potential Labor Force}) \times 100 / (\text{Labor Force} + \text{Potential Labor Force})$

The ICLS framework can be applied to that of the NSSO in the following manner. While LU1 would continue to be the definition of the UR as used by the NSSO thus far, the LU2 measure would need to be adapted to NSSO's

**FIGURE 1. New Measures of Labor Underutilization to Complement the Unemployment Rate**



Source: International Labour Organization (2018).

concepts. We propose creating a modified measure of underemployment for India, which includes visible underemployment as defined by the NSSO above. Ideally, we would like to incorporate a measure of invisible underemployment in this framework but given that there is no single metric of this measure, including it in the definition of LU2 is tricky. Nevertheless, it is important that we continue to separately report the three measures of invisible underemployment that the NSSO has reported in the 68<sup>th</sup> Round, that is, the usual Principal Status workers who did not work regularly throughout the year; the ratio of persons who sought or were available for additional work in the US; and the ratio of persons who sought or were available for alternative work.

To construct LU3, we would need to have a measure of the potential labor force, a concept that has not been explicitly laid out in household surveys in India thus far. The Resolution adopted at the 19<sup>th</sup> ICLS defined the potential labor force as all persons of working age who, during the short reference period, were neither in employment nor in unemployment, and:

- (a) Carried out activities to “seek employment,” were not “currently available,” but would become available within a short subsequent period established in the light of national circumstances (i.e., unavailable job seekers); or

- (b) Did not carry out activities to “seek employment,” but wanted employment and were “currently available” (i.e., available potential job seekers).

The potential labor force can be computed from the information in Block 6 of the NSS schedule (Table 5). The workers defined in Category (b) above can be computed from the information in Columns 7–9 of Block 6. In order to estimate those in Category (a) above, we would need to modify the question in Column 9, that is, “whether made any effort to get work” to not just those who reported codes 1 (i.e., sought/available work on most days) and 2 (i.e., sought/available work on some days) in Column 8, but also those who reported code 3 (i.e., who were not available). Additionally, the difference between seeking/not seeking work and being available/not available for work needs to be spelt out explicitly to capture the potential labor force.

The idea of measuring the potential labor force is relevant in the Indian context. With a rise in education levels and income levels, the aspirations of the educated youth have risen. They may no longer be willing to join the labor force if they feel that they will be unable to find a suitable job commensurate with their qualifications or desires. In this context, it would be useful to separately identify a category of discouraged job seekers in line with the recommendations of the 19<sup>th</sup> ICLS. This category comprises those workers who did not “seek employment” for labor market related reasons such as their past failure to find a suitable job, lack of experience, qualifications or jobs matching the person’s skills, lack of jobs in the area, and being considered too young or too old by prospective employers.

Creating the above-mentioned four headline labor underutilization indicators will enable us to have a wider picture of the extent of labor underutilization. In order to understand where the challenges lie from a policy perspective, it would also be useful to look at the composition of labor underutilization. This can be done by calculating the shares of unemployment, underemployment (both visible and invisible), and the potential labor force in overall labor underutilization.

### *4.3. Including Additional Activity Status Classifications*

Household surveys report various activity status classifications for the employed—self-employed, casual labor, and regular wage salaried employees. However, over time, the heterogeneity of employment types has increased, and new forms of standard and non-standard employment have emerged. Thus, there is a need to rethink the existing classification and incorporate additional categories of employment. Although we are not

proposing doing away with the existing status classification as that would introduce problems of comparability over time, it would be useful to add additional sets of questions in the survey to understand changes in the nature of employment and the relationship between employers and employees.

This can best be explained through the example of homeworkers who are defined as those who carry out work for remuneration in premises of their choice, other than the workplace of the employer, resulting in a product or service as specified by the employer, irrespective of who provided the equipment, material, or inputs used (ILO Home Work Convention, 1996). Homeworkers constitute a subcategory of home-based workers (i.e., those who carry out remunerative work within their homes). Traditionally, NSSO has considered the homeworker or outworkers as self-employed in its surveys. However, in the NSS 55<sup>th</sup> Round, they categorized these workers separately by posing a set of additional questions about the location of their workplace, whether they worked under given specifications, whether they were provided equipment, raw material, or inputs by the employer, and if they received remuneration on a piece rate basis. These questions have, however, been discontinued after the NSS 61<sup>st</sup> Round, and do not find mention in the PLFS. The identification of homeworkers or outworkers from among the self-employed is crucial given that this category accounts for half the workforce, and we need a better understanding of the nature and quality of self-employment.

The category of RWS employment also merits inclusion of different types of contracts. In addition to permanent jobs, the RWS classification includes several temporary jobs. The nature of contracts is changing as short-term and temporary jobs are becoming more prevalent. For instance, in 2018, the government introduced the provision of fixed-term employment to provide enterprises flexibility to adjust their workforce. Thus, there is scope for suitably restructuring the codes for types of job contracts for RWS workers. Additionally, the last two decades have seen a surge in contract worker usage. While the enterprise surveys make a distinction between workers directly hired by enterprises and those hired via contractors, the classification used in household surveys does not allow us to distinguish contract workers from others. Typically, these workers are subsumed in the category of RWS workers. Even in the enterprise surveys, there is very limited information about contract workers. The ASI schedule does not even collect information on the gender breakdown of contract workers.

Finally, another classification of workers, which is attracting attention globally, is that of gig economy workers. Typically gig workers have non-traditional contracts and work as independent contractors or freelancers. They are included in the count of the self-employed. However, as the

gig economy emerges as an important avenue of employment generation, it would be useful to measure and understand trends in the growth of such non-traditional working arrangements. As there is no internationally accepted definition of gig economy workers, it would be difficult to create a separate category or code for them in the existing activity status classification. However, given that a key feature of the gig economy is the use of digital labor platforms, it might be useful to design and incorporate structured questions around defining features of the gig-economy work to identify workers. As mentioned above, it could be challenging to incorporate such detailed questions in the annual PLFS. They could instead be posed in the quinquennial EUS.

#### *4.4. Estimating Informal Employment and Understanding Conditions of Employment of Workers*

As described in Section 2, India's labor markets are characterized by informality. Given the significance of the informal workforce and the need to understand the quality of employment, it would be relevant to include the "rate of informal employment" as a KILM and report this measure in all official reports when survey data is released. However, computing such a statistic is fraught with challenges unless we have a clear definition of what constitutes informal employment.

The 15<sup>th</sup> ICLS in 1993 defined the informal sector as consisting of all private unincorporated enterprises (excluding quasi-corporations), that is, enterprises owned by individuals or households that are not constituted as separate legal entities independently of their owners, and for which no complete accounts are available that would permit a financial separation of the production activities of the enterprise from the other activities of its owner(s).

However, this definition was criticized as it captured only work that took place in unincorporated enterprises. The notion of informal employment is much wider as it includes several additional types of employment outside informal enterprises, in particular, persons working in formal enterprises who are not covered by social protection through their work. Consequently, in the 17<sup>th</sup> ICLS in 2003, a broader concept of informal employment comprising the total number of informal jobs, whether carried out in formal sector enterprises, informal sector enterprises, or households during a given reference period, was laid out.

The two-dimensional matrix in Figure 2 provides a conceptual framework for estimating formal and informal employment as per the 17<sup>th</sup> ICLS. The framework uses a building block approach which disaggregates total

**FIGURE 2. Conceptual Framework for Estimating Informal Employment**

Production Units by Type	Jobs by Status in Employment									
	Own-account Workers		Employers		Contributing Family Workers	Employees		Members of Producers' Cooperatives		
	Informal	Formal	Informal	Formal	Informal	Informal	Formal	Informal	Formal	
Formal Sector Enterprises					1	2				
Informal Sector Enterprises (a)	3		4		5	6	7	8		
Households (b)	9					10				

Informal jobs    
  Formal jobs    
  No jobs by definition

Source: Adapted from Bhalla (2009), citing Hussmanns (2004).  
 Notes: (a) As defined by the 15<sup>th</sup> ICLS (excluding households employing paid domestic workers).  
 (b) Households producing goods exclusively for their own final use and households employing paid domestic workers.  
 Cells in dark gray refer to jobs, which by definition do not exist in that type of production unit. Cells shaded in light gray refer to formal jobs (including cell 7 in informal sector enterprises). Unshaded cells refer to different types of informal jobs. Thus, *Informal Employment* is comprised of Cells 1–6 and 8–10. *Employment in the informal sector* is represented by Cells 3–8. *Informal employment outside the informal sector* is represented by Cells 1, 2, 9, and 10.

employment according to two dimensions: type of production unit and type of job. The type of production unit is defined in terms of legal organization and other enterprise-related characteristics, while the type of job is defined in terms of status in employment and other job-related characteristics.

In the Indian context, it would be useful to construct a measure of “informal employment” as defined in the above matrix, which combines the enterprise-based and jobs-based definitions of informality. This is particularly pertinent in the light of the increasing informalization of the formal sector. Reporting such a combined measure along with the UR would enable us to understand the overall changes in the quality of employment.

It is worth pointing out that the ICLS has not endorsed the term “employment in the informal economy,” which has been used by the ILO to refer to the sum of employment in the informal sector and informal employment outside the informal sector. The 17<sup>th</sup> ICLS noted that for statistical purposes,

it would be better to keep the concepts of informal sector and informal employment separate. This was perhaps a result of the fact that different operational criteria were being used for defining informal jobs in different countries and a combined metric would not be comparable across countries. However, for the purpose of getting a comprehensive measure of informality in India and understanding what is happening to the quality of employment over time, this would be a useful exercise.

#### *4.5. Extending the Quarterly Module to Rural Areas*

A significant feature of the PLFS has been the introduction of a quarterly module in urban areas which will allow us to generate KILM at quarterly frequency. However, often the workforce that resides in rural areas commutes to urban areas to work, and vice versa. For instance, Chandrashekhar (2011) estimates that a total of 12.42 million non-agricultural workers commuted across the rural–urban boundary, in one direction or the other, for work in 2009–10. Labor force surveys estimate the size of the workforce by place of residence and not by workforce. Typically, the size of the rural (urban) workforce is set equal to the number of workers living in rural (urban) areas. Given the significant movement across rural and urban areas, there is a need to make adjustments in estimates of the rural and urban workforce accounting for the fact that several rural residents work in urban areas, and vice versa (Mohanani 2008).

The challenge of addressing the commuting worker becomes a serious cause of concern now that the rural and urban modules will be held at a different frequency. It would be appropriate to extend the quarterly module to rural areas as well and make the necessary adjustments for commuting workers. This will also enable us to understand the size of the rural–urban linkage. However, it needs to be noted that the exercise of extending the quarterly module to rural areas will be meaningful and worthwhile only if the sample size is large enough that the quarterly changes can be measured with precision. In fact, at this point, it is unclear if the sample size in the urban quarterly module is also large enough to accurately measure changes in urban labor market parameters.

#### *4.6. The Need for a Business Register*

While we have two separate establishment surveys for the organized and unorganized sectors, in practice, the distinction between the two is not so neat and clean. In principle, the unorganized manufacturing sector should not include units that use electricity and hire more than 10 workers, and those that do not use electricity but nevertheless employ 20 or more workers.

Nevertheless, it is often noted that many big units with a sufficiently large number of workers are included in the NSS Unincorporated Enterprise Survey. For instance, in 2015–16, over 12,000 of the surveyed enterprises in the NSS 73<sup>rd</sup> Round hired 10 or more workers. In the 2010–11 survey, there were close to 10,000 surveyed units having 10 or more workers (Kapoor 2017). These units should have, in fact, been in the ASI frame. Just as there are several larger units included in the NSS frame, it has been observed that there are several units hiring less than 10 workers which have been reported in the ASI database. The chief inspector of factories in each state maintains the live register of factories that form the frame for conducting the ASI. However, it is often observed that the live register includes several factories which have been closed for years. Additionally, there is large-scale evasion of registration under the Factories Act (Nagaraj 2002). The above-mentioned exclusions and inclusions have serious implications for estimating not just employment but also GVA and GVA per worker in both the organized and unorganized sectors.

Given the above problems, there is an undisputed need to get the ASI frame right. With this objective, the National Statistical Commission (NSC; 2001) recommended the creation of a “Business Register,” a list which would cover all enterprises with at least 10 workers, including those covered under ASI and public sector enterprises. It was suggested that the creation of the Business Register would be associated with the development of a unique coding system (or a Business Identification Number) identifying all enterprises included in the list. The use of the unique code by the units would be made mandatory for purposes like paying sales tax, electricity bill, telephone bill, or in getting facilities like bank loans.

Despite this recommendation, the idea of the Business Register has hardly taken off. According to NSSO’s recently released Technical Report on the Services Sector (2019), only 11 states have implemented it.<sup>12</sup> And here, too, their coverage is skewed and limited.<sup>13</sup> While the Department of Company Affairs has initiated a computerized 21-digit code for registered companies, called the Corporate Identity Number, this alone is not enough. The idea needs to be extended to a unique business numbering system that is mandatory for all factories/companies, and for use in all official transactions (Nagaraj 2002). Such a numbering would be like a social security number system used in the developed countries for their citizens. MoSPI has

12. This includes Andhra Pradesh, Arunachal Pradesh, Himachal Pradesh, Manipur, Nagaland, Punjab, Rajasthan, Tamil Nadu, Telangana, Uttar Pradesh, and Uttarakhand.

13. In the Business Register, 7.1 percent of all sample enterprises were non-traceable/closed and 13.5 percent were misclassified/out of coverage.

indicated that it plans to compile a national business register of all business enterprises in the country that would be updated periodically, but there has not been much progress on this front.

An additional problem vis-à-vis enterprise surveys is the under-listing of enterprises in the EC. The above-mentioned enterprise surveys are, in fact, follow-up surveys of the EC and the inability to get an accurate count of enterprises in the EC creates problems in the follow-up surveys, too. It has been observed that the number of enterprises and the number of workers reported in the EC is, in fact, lower than the Unincorporated Enterprise Survey estimates. Manna (2010) explains that one of the limitations of the EC is that the number of own-account manufacturing enterprises (OAMEs) estimated from it are significantly lower than those reported by the NSS Enterprise Survey. He finds that the total number of establishments in the EC (2005–06) is about 22 percent lower than NSS's Survey of Unorganized Manufacturing Enterprises conducted during the same year. For the most recently conducted EC and NSS Survey of Unincorporated Non-agricultural Enterprises (2015–16), we once again find this discrepancy. In the manufacturing sector specifically, the EC reports that the total number of enterprises without hired workers is 7.21 million. For the NSS Enterprise Survey, however, we find the number of OAMEs to be 16.8 million. But this figure should, in fact, be lower than the EC's figure given that the latter must have universal coverage. Manna and Bhattacharjee (2004) have noted that the under-listing is largely confined to smaller or invisible units or to those units carrying out their activities without fixed premises.

The importance of ensuring universal coverage in the EC cannot be emphasized enough. This is essential to have a single accurate frame that captures non-agricultural enterprises. This requires additional resources to be spent on the survey and the use of modern technology. In fact, it is largely because of the absence of an appropriate frame for enterprises that household surveys have been more widely used for data collection on employment statistics. There is an acute need to strengthen sources of enterprise and establishment level data. This requires wider coverage across enterprises of different sizes and sectors. It also requires enterprise surveys, in particular the unincorporated enterprise survey, to be carried out at a greater frequency. The EC should also be conducted at regular intervals.

#### *4.7. The Role of States*

A key feature of the process of data collection undertaken by the NSSO is that for each round, two types of samples known as "Central samples" and "state samples" are allotted for conducting these surveys. The surveys of

Central samples are conducted by the Government of India, while those of state samples are conducted by state agencies. The sizes of Central and state samples are equal for most of the states/UTs (equal matching sample). In some states, the number of sample units surveyed by state statistical agencies is double that of the size of the Central sample. One of the objectives of the states' participation in the NSS program is to provide a mechanism by which the sample size would increase and the pooling of the two sets of data would enable better estimates at the lower sub-state level, particularly at the district level. At the state level, this will increase the precision of the results.

Given the importance of pooling the Central and state sample data, the 13<sup>th</sup> Finance Commission made a special provision for additional funds in each district to carry out this exercise. Despite this, little progress was made in terms of evolving a uniform methodology of pooling and testing for poolability of the two sets of data. Only a few states pooled the results of the Central and state samples for the NSS Employment–Unemployment rounds. And here too, there was often a lack of uniformity in their approach. This resulted in a loss of comparability of pooled data. It was against this backdrop that the NSC appointed a professional committee under the chairmanship of Dr R. Radhakrishna, former Chairman, NSC, to examine the above-mentioned issues. The Committee in its report gave a detailed methodology for pooling and suggested tests for poolability.<sup>14</sup> Following the recommendations of the Committee, the Data Processing Division of the NSSO took initiatives to provide technical guidance and support to the states in pooling the data collected in the NSS 66<sup>th</sup> Round.

However, despite all these efforts, only a few states have made attempts to process and tabulate the state samples such that they could be merged with the Central samples to generate more reliable state-level estimates and, more importantly, district-level estimates. The absence of any district-level employment data barring the economic tables provided in the Population Census described in Section 2 is a problem and points to the importance of pooling the state and Central samples. In fact, in the PLFS, standard errors at the state level are relatively large, making these results less reliable. Involving state agencies in data collection is indeed imperative. This necessitates greater initiative on the part of the states to invest in data collection

14. It suggested that the poolability test of two sets of data must be exercised before pooling the two independent estimates derived at a particular domain using weights as an inverse of estimated variances of the estimates or using weights as a matching ratio of the States' participation in the NSS program. Thus, in contrast to earlier methodologies of pooling the data by merging two datasets and recomputing weights based on merged data, the new methodology suggested the pooling of estimates using weights as stated above.

and establish a mechanism of rigorous data quality checking through inspection, scrutiny checks, training, and managing the field operations in a time-bound manner.

#### *4.8. Presentation of Survey Results in a User-friendly Manner*

It is often noted by data users that the results of the surveys, in particular, household surveys, are not presented in a readily usable form. This can be elucidated through the following example. If one wanted to know from the survey report the size of the labor force and workforce at the national level, this would not be a simple calculation. First, the report lays out the labor force participation and worker population rates separately for rural male, rural female, urban male, and urban female, for the three different concepts (US, CWS, and CDS). Second, these rates would need to be multiplied with the population figures obtained from the Census and not the population estimates obtained from the household survey. This is because it has been observed that the population estimates derived from the NSSO are, in general, on the lower side as compared to the Census population. This difference arises mainly due to the differences in coverage and methods adopted in NSS in comparison with the Census operation. While a note of caution is always sounded in the report on this issue, it often creates considerable confusion among users. To avoid such confusion, it would be useful if the NSSO provides an additional summary publication giving the estimates of KILM (as well as informal employment and the extent of labor underutilization) using the Census population. Also, to make these estimates internationally comparable, it would be appropriate to report results based on the reference period of a week as is done in most other national labor force surveys.

#### *4.9. Leveraging Technology to Conduct Surveys*

Given the vastness of the country and the varied sources from which data are to be compiled, transmitted, tabulated, and analyzed, a manual system of data compilation and transmission is hardly viable in India. There is a need for massive computerization of labor statistics with computer networking so that online data are available at specified points, the delay in transmitting data from the field is minimized, and dissemination of data is speedier. All this is easily possible given the current state of information technology in the country.

The PLFS has taken a major leap forward in this context by collecting information from sample households using the CAPI method. Prior to the PLFS, data under the socio-economic surveys of NSSO were collected from

the field by using paper schedules. After the completion of data collection, it took about a year to make available the results of the survey. Given this delay, the NSS adopted the World Bank-CAPI solution platform for the PLFS. Data for the PLFS were collected in the field using tablets and the CAPI solution with an in-built data validation process. This not only proved useful in the collection of primary data from the households but also helped reduce the time involved in data transfer and processing of the survey results. CMIE's experience also shows how data collected on a specially loaded application on mobile phones can be validated and uploaded for use by the end of day, every day, thereby enabling generation of high-frequency estimates. The forthcoming EC (2019) also proposes to include geo-tagging data of business enterprises to ensure that no bogus enterprises are included in the official statistics. This will help clean the data organically as field officials will enter the live location of enterprises via mobile phones. Thus, there is considerable scope for using modern technology to improve the accuracy and reliability of data and reduce the time spent in data collection, verification, processing, and dissemination by leveraging technology. To exploit these potential gains, a concerted effort is required not just to increase investments in technology but also to train investigators in the use of these tools.

## 5. Conclusion

As countries the world over are pressed to produce labor market statistics at quarterly or monthly frequency, the desire to produce high-frequency employment data in India is understandable. However, simply aggregating the number of unemployed and employed at a macro level periodically serves little purpose in the Indian context. Employment statistics are not very meaningful unless they are able to convey information about the quality of jobs being created and how the conditions of employment are evolving.

Following the recommendations of the 2017 Task Force set up to revamp the employment data architecture, India's employment statistics are undergoing significant revision. Impatience with the absence of high-frequency data has led to a frantic quest for such data. This search has culminated in the recent release of monthly estimates of administrative data obtained from the EPFO, ESIC, and NPS. Despite several caveats, it is disconcerting that the EPFO database, which has never been made available in the public domain, is being viewed as a substitute for NSSO's household surveys. The EPFO databases are still at a very nascent stage and need to be cleaned up and processed to produce estimates that can be of any significant use. In the

best of circumstances, administrative data can only give us a partial estimate of formal employment, let alone be a replacement for household surveys.

NSSO's EUS, which is a very elaborate and comprehensive survey, has done a remarkable job of capturing the complex realities of India's labor market and provides a rich source of data for decades. The importance of conducting NSSO's EUS, which asks in-depth comprehensive questions, cannot be emphasized enough. Of course, given its level of detail and costly operations, it cannot be conducted annually. Therefore, it should be conducted quinquennially and supplemented with the PLFS.

While many have argued that India is in need of a fresh employment data architecture, we believe that the focus should not be on conducting new surveys. India's statistical data collection machinery, in particular, household surveys, have been lauded globally. Instead of dispensing with the current structure, it is worth trying to understand how the existing machinery can be strengthened to produce more relevant and comprehensive labor market data. In the foregoing discussion, we have tried to illustrate relatively better ways of understanding changes in the employment landscape. To address the information gaps, we propose asking more probing questions about the nature of the employment arrangement and contracts, extending the quarterly module of the PLFS in rural areas, and putting in place a National Business Register, which provides a complete frame for follow-up enterprise surveys. Given that it is the unemployment rate that invariably attracts the most attention in the policy discourse each time survey results are released, we propose creating estimates of informal employment and labor underutilization and reporting these with NSSO's other Key Indicators of the Labour Market. This will be particularly important for understanding the gravity and enormity of India's employment challenge.

The discussion in this paper in no way claims to be an exhaustive coverage of employment data issues in India. In addition to the issues described here, significant challenges loom. Some key concerns deserve mention here. The first is the absence of a good dataset on the services sector. The second is the lack of understanding about the activity pattern of those who are outside the labor force. This is particularly important in a young economy where the youth inactivity rate (i.e., the share of youth who are not in employment, education, or training) stands at over 20 percent. Third, qualification and skill mismatches are becoming a critical issue of policy concern as the world of work is being transformed by the forces of technology, demographics, globalization, migration, and climate change. The absence of data on workers' skills and the task content of occupations poses a serious challenge in designing skilling policies in India. In fact,

establishment surveys in India collect no information on the education or skill qualification of their workers.

Implementing any of the suggestions laid out in this paper requires strengthening the capacity of the national statistics systems. This entails not only greater investments in the data collection machinery but also addressing the shortage of regular field investigators faced by the NSSO and training them to ensure that they can correctly implement the concepts and definitions applied in the survey. However, against the backdrop of the controversy surrounding the results of the PLFS (2017–18) Report, what is even more significant is to restore faith in India’s statistics and re-establish the credibility and independence of the institutions that produce them. Unless we recognize official data as a “public good” (Mohanan 2019b) and provide researchers access to data to enable informed debates and discussions, any effort to improve the employment data architecture would serve little purpose.

## References

- Bhalla, S. 2009. “Definitional and Statistical Issues Relating to Workers in Informal Employment.” *Working Paper No. 3*. New Delhi: National Commission for Enterprises in the Unorganized Sector, January.
- Chandrasekhar, S. 2011. *Estimates of Workers Commuting from Rural to Urban and Urban to Rural India: A Note*. Mumbai: Indira Gandhi Institute of Development Research.
- Dantwala Committee. 1970. “Report of the Committee of Experts on Unemployment Statistics.” New Delhi: Planning Commission of India, Government of India.
- Ghose, A.K. 2016. *India Employment Report 2016: Challenges and the Imperative of Manufacturing-Led Growth*. New Delhi: Institute for Human Development.
- Hussmanns, Ralf, Farhad Mehran, and Vijay Verma. 1990. *Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods*. Geneva: International Labour Office.
- Hussmanns, Ralf. 2004. “Measuring the Informal Economy: From Employment in the Informal Sector to Informal Employment,” *Working Paper No. 53*, Policy Integration Department, Bureau of Statistics. Geneva: International Labour Office.
- International Labour Organization. 2018. *Avoiding Unemployment Is Not Enough: An Analysis of Other Forms of Labour Underutilization*. Geneva: International Labour Organization.
- Kapoor, R. 2017. “Waiting for Jobs.” *Working Paper No. 348*. New Delhi: Indian Council for Research on International Economic Relations, November.
- . 2018. “Rethinking India’s Employment Data Architecture.” *Economic & Political Weekly*, 53(40): 14–16.

- Manna, G.C. 2010. "Current Status of Industrial Statistics in India: Strengths and Weaknesses." *Economic & Political Weekly*, 45(46): 67–76.
- Manna, G.C. and J.P. Bhattacharjee. 2004. "Enlarging the Scope of Annual Survey of Industries." *Economic & Political Weekly*, 39(48): 5163–5166.
- Ministry of Statistics and Programme Implementation. 2019. *PLFS Annual Report 2017–18*. New Delhi: Ministry of Statistics and Programme Implementation, Government of India.
- Mohanan, P.C. 2008. "Differentials in the Rural-Urban Movement of Workers." *The Journal of Income and Wealth*, 30(1): 59–67.
- . 2009. "Background Paper for the Workshop on Conceptual Issues in Measurement of Employment–Unemployment." Available at [http://mospi.nic.in/sites/default/files/workshop/nsc\\_background\\_paper\\_12jan09.pdf](http://mospi.nic.in/sites/default/files/workshop/nsc_background_paper_12jan09.pdf) (accessed May 28, 2021).
- . 2019a. "Data on Employment–Unemployment: An Appraisal", Unpublished Paper.
- . 2019b, February 12. "Because Data Is a Public Good." *The Indian Express*. Available at <https://indianexpress.com/article/opinion/columns/because-data-is-a-public-good-nsc-p-c-mohanan-employment-survey-5578946/> (accessed 28 June 2019).
- Nagaraj, R. 2002. "How to Improve India's Industrial Statistics." *Economic & Political Weekly*, 37(10): 966–970.
- National Sample Survey Office. 2015. *Employment & Unemployment Situation in India, NSS Report No. 563, 68<sup>th</sup> Round, July 2011–June 2012*. New Delhi: Ministry of Statistics and Programme Implementation, Government of India.
- National Statistical Commission. 2001. *Report of Dr Rangarajan Commission*. New Delhi: Ministry of Statistics and Programme Implementation, Government of India.
- NITI Aayog. 2017. *Report of the Task Force on Improving Employment Data*. New Delhi: NITI Aayog, Government of India.
- Papola, T.S. 2014. *An Assessment of the Labour Statistics System in India*. New Delhi: International Labour Organization Country Office for India.
- Planning Commission of India. 2002. *Report of the Special Group on Targeting Ten Million Employment Opportunities per Year over the Tenth Plan Period*. New Delhi: Government of India.
- Work Statistics Committee. 2013. "Report II: Draft Resolution Concerning Work Statistics." Paper presented at International Conference of Labour Statisticians (ICLS), October 2–11. Geneva: International Labour Office.

To view the entire video of this IPF session and the General Discussion that ended the session, please scan this QR code or use the following URL:

<https://www.youtube.com/watch?v=4dZTQd6fvMA>



# Comments and Discussion\*

**Chair: Pravin Srivastava**

*Secretary, Ministry of Statistics and Programme Implementation  
Government of India*

## **G. C. Manna**

*NCAER and Former Director General, NSSO and CSO*

The author gives a clear description of various data sources for employment (and unemployment) statistics in India, along with coverage of the respective data sources, and strengths as well as weaknesses of the datasets emanating from them. As suggested by the author, India's labor market has a dualistic structure, which is characterized by the presence of an organized sector that co-exists with a large unorganized sector. This fact is substantiated in the paper by citing the related estimates (Table 3 of the paper) of the number of workers by organized–unorganized enterprises and formal–informal employment based on the NSS 61<sup>st</sup> Round (2004–05) and NSS 68<sup>th</sup> Round (2011–12). As regards the estimates of the aggregate number of workers presented in this table, the author needs to clarify how the aggregates have been derived and for doing so, whether the NSS employment rates have been multiplied with the projected population based on the Census, and if so, the reference date considered for the projected population. It would also be useful to clearly spell out the methods followed in categorizing the workers under the organized–unorganized and formal–informal break-ups. Further, the sector-wise classification of workers in which services is shown as a separate sector beyond non-manufacturing needs modification because the services sector comes within the ambit of the non-manufacturing sector as per the standard concept.

Under “Issues and Recommendations” (Section 4), the paper describes the challenges and difficulties in interpreting the CWS-based estimates of employment, as per the existing definition, by also narrating the views of some other authors in this regard. The author finally advocates the need

\* To preserve the sense of the discussions at the India Policy Forum, these discussants' comments reflect the views expressed at the IPF and do not necessarily take into account revisions to the conference version of the paper in response to these and other comments in preparing the final, revised version published in this volume. The original conference version of the paper is available on [www.ncaer.org](http://www.ncaer.org).

for continuation of all the three approaches of measurement of labor force parameters, viz., Usual Status, Current Weekly Status, and Current Daily Status. Given the complexities of India's labor market situation with a huge percentage of workers employed in the informal sector and in short-term employment, this recommendation seems to be a reasonable one. As regards the suggestion of some authors to change the definition that has been cited in the paper, it may be useful if the author also highlights the effect of such a change in the cross-country comparison of CWS-based estimates, since many countries may be following the same definition as adopted in the NSS.

Based on the existing approach of data collection, the problem of estimating the total number of jobs has been discussed in Section 4.3, citing, for example, the difficulty in capturing multiple activities based on the CDS approach given the provision to record at the most two activities on a single day. In this context, it may be useful for the author to highlight the extent of the prevalence of more than two activities by analyzing the information on day-wise total hours actually worked, while considering all the work vis-à-vis the time spent in maximum two activities performed during the day that is being collected in the ongoing PLFS. The difference between the two times as stated above is likely to throw some light on the prevalence of multiple activities.

Regarding the suggestion (Section 4.4) to incorporate additional categories of status classification to capture new forms of standard and non-standard employment and restructure the codes for types of contract, one needs to be a little careful because the modification of the existing code structure might lead to data comparability issues. Instead, it may be more appropriate to include a supplementary set of questions to meet the proposed additional data requirements, without disturbing the existing classification.

In Section 4.6, the author recommends extension of the quarterly module of the PLFS to rural areas as well. Before such a step is taken, it would be useful to study the sample size requirements since my limited data analysis and study in this regard suggest the need for a drastic increase in the sample size if one has to detect changes in the quarterly labor force parameters at the state level through the PLFS.

After highlighting the limitations of the EC database and the inadequacies of the ASI frame in terms of erroneous exclusions (of units with 10 or more workers) and inclusions (units with less than 10 workers), the paper underlines the need to implement a National Business Register (Section 4.7) that could serve as a frame for various establishment surveys. The cited problem of imperfect frames is a genuine one and concerted efforts are undoubtedly warranted for developing a proper sampling frame. Similar

to the evidence given of a number of larger manufacturing units that got covered in the NSS 73<sup>rd</sup> Round, though they were supposed to be covered in the ASI frame instead, a similar study based on the NSS 73<sup>rd</sup> Round could also be carried out for the services sector (including trade) to throw some light on the weakness, if any, of the EC frame of services sector establishments with 10 or more workers.

Finally, the recommendation to continue the traditional NSS quinquennial round along with the annual PLFS appears to be a reasonable one, given the importance of additional items of information, relevant in the context of measurement of labor underutilization, included in the quinquennial round but dropped from the PLFS module.

## **Rinku Murgai**

*World Bank*

My task is easier because we have seen a very clear presentation and excellent comments as well from Dr Manna. I won't use up time summarizing what has been said. I agree with many of the things that Radhicka says in her paper. I would only suggest that for anybody who is going to work on employment data in India, this is going to become a must-see reference manual. It is very comprehensive, clearly done, and excellent.

Let me focus on the big picture questions here. The first question I was asking while reading the paper was: Is India's employment data architecture in good shape, and where have we come? I think there has been a lot of concern. The last comprehensive round until very recently was 2011–12, and there were various sporadic efforts, which were hard to interpret—data being released, data not being released—and therefore lots of reasons for concern. The fact that the PLFS is now in place is a major advancement. I think we have taken a big step there. There are reduced time delays in processing due to the adoption of CAPI, which is a big step. There is a new dissemination policy issued by the MoSPI. If adhered to, that is a major step in the right direction. There is adoption of the International Household Survey Network approach for the systematic production of metadata. We are seeing some important steps that are relevant not just for the employment data architecture but for other data as well.

But there is obviously much more to be done. Anybody who has not been living in a bunker knows that there is a lot of noise around India's statistics. So the question is: What are the priorities and how should the priorities be set? Let me talk about four broad areas. One of them concerns the set of

priorities in improving the employment data that comes from household surveys. There I was going to endorse the suggestion of a quarterly PLFS in rural areas as well, but I agree with Dr Manna's point that it has to be done carefully in relation to the resources that are available and can be mobilized by the Ministry to expand sample sizes. Clearly, it seems that we also need to potentially expand sample sizes in urban areas if we are going to detect changes, but I think the quarterly addition to the annual survey is an important step that needs to be taken.

Second, I also endorse the idea of supplemental modules, not necessarily reviving the EUS quinquennially but concurrently every few years with the PLFS. In the areas where PLFS is being done, supplemental modules may be added. Many countries do this on questions that do not require annual or quarterly monitoring. Those suggestions made a lot of sense to me.

The third question, which I think is in the paper but to my mind is not done sufficiently and systematically, is experimentation. There are questions, for example, when you look at what has been happening to India's female labor force participation. When you compare that with the population Census, and with the NCAER India Human Development Survey, in some cases you get different trends. So the question is, why that is so. Is it something about the questionnaire design, is it something about the way the surveys are being fielded? We need to get a much better understanding of these issues, and this is something that the Ministry must do. It could be in partnership with think tanks, such as MoSPI's partnership through its MoU with NCAER, it could be with other agencies. MoSPI and NSSO have been used to doing experiments. The 55<sup>th</sup> Round created a lot of noise, but it came at the end of a lot of experiments that were very useful. It is important to do those experiments more often, more systematically.

In this context, the suggestion to extend to the gig economy by adding a set of questions has to be done in experimental mode so that we can understand what works and what does not. There is a time-use survey that MoSPI is going to do. If properly designed as an experiment and analyzed, it will provide insights into how you might improve the measurement of women's work.

Here, I have a model in mind that Statistics Canada follows—they have what they call “pathfinder projects.” When there is an issue that is clearly of relevance to citizens, to the public, to policymakers, they mobilize teams and resources to specifically look for solutions on how they can strengthen both the data collection mechanisms and analysis. The advantage of signaling that these are pathfinder projects is that it avoids the confusion that happened when the analysis of the payroll data was released in India. As Radhicka says, what is being done is somewhat experimental. But I think it must be

done. We need to move in that direction not as a substitute for household surveys but for supplementing our understanding of what is going on in the economy. However, it is important to put them in the right context, that these are experimental, so that they are not misquoted and misunderstood by both the users and policymakers. It is the job of the statistics agency to signal that they are in experimental mode in some things before a decision is taken to adopt and scale up.

The fourth area that I would suggest on the household survey side is the need for conceptual clarity. There are many examples in Radhicka's paper where she has already clarified several concepts. With the PLFS becoming a major source of employment data, going forward, there has to be conceptual clarity on comparability of the PLFS with previous surveys. I am not a statistician, but my understanding of the stratification design is that the headline numbers at the all-India level and at the state level are indeed comparable to the EUS. That is not what the official position of the government is, but I think this is the kind of debate that is needed so that clarity may emerge.

There are other reasons why the PLFS may or may not be comparable to previous rounds. It is a panel survey. We need to understand attrition. There are issues related to non-compliance rates that seem to be different from previous rounds. The PLFS was fielded using CAPI, not paper. Again, it is important to get the research in place that allows you to assess whether or not using the CAPI approach versus the paper-based schedule has changed the responses. It has to be a routine exercise but should provide clarity, and it requires both the statistics agencies and the research community to work on these questions.

The final point I would make on the household survey side, which I think will add to the data architecture, is to have an announced advance release calendar. Just the way it exists for parts of the national accounts in compliance with the Special Data Dissemination Standard (SDDS), an advance release calendar will help add trust and credibility to the system. This is already stated as a proposal in the new dissemination policy.

On the enterprise survey side, let me just make one comment. I think there the priority must be to build an integrated statistical business register. It will help not only with the frame but also with moving towards getting much more real-time data on economic demography, not on the informal sector which is where the vast numbers of people are employed, but on the formal sector. The opportunities now are tremendous. With the Economic Census (EC) going into the field very shortly, the availability of the Goods and Services Tax Network (GSTN) and the Ministry of Corporate Affairs (MCA), it is very important to launch a research program that starts looking

at the differences between these different sources, the opportunities for linking the EC with the GSTN and MCA. It means working very differently for MoSPI. It means signing memorandums of understanding to access the GSTN and MCA data. The kind of careful work that was done for the 74<sup>th</sup> Round is what is needed, in addition to other building blocks, for developing a business register.

I am less pessimistic on the use of administrative data for starting to track real-time measures of changes in firms, and to my mind there are three issues. One is that we have to get the foundational statistical infrastructure in place. Currently, data is extremely siloed and the classifications, issues of consistency over time, and many other such issues make it very difficult to actually use administrative data. There needs to be a strong data quality assurance framework. There is a framework that has been notified, not yet operationalized—that has to be a big priority. There are questions about whether or not there is a need for a chief data officer in India, and how to handle that in a decentralized statistical system.

My last point pertaining to the employment data strategy for India, to stronger national accounts, and to strengthening any arm of the national statistical system, is that we have to ask the question whether our institutions are strong enough, whether we have a legal framework that permits data aggregation and sharing of databases, and whether rules are in place that enable data flows under confidentiality guarantees. I would also say it is time to think about whether the right roles are currently assigned between the Chief Statistician of India and the NSC. Is there too much overlap? Are both entities sufficiently empowered? My own feeling is that it is time to do a white paper around these questions and then figure out whether we should go for legislative backing. For instance, as regards the NSC, currently it is not a statutory body, but it was meant to become a body that had legislative backing through the NSC Bill. I know this is a little far removed from the issues that Radhicka talks about in her paper, but they are fundamental to having any data architecture that is strong enough to deliver what users need, that is, a system that is credible, and is seen to be credible.

## General Discussion

Sudipto Mundle concurred with the paper that conceptually household surveys and employment and enterprise surveys would not be comparable, but at an aggregate level, they should be and should move in the same direction. He asked what might explain the huge difference seen in the paper.

Kaushik Basu asked about labor force participation rates, a major concern in India, and whether it was possible to say anything about what “unavailable” means. The paper touched upon ILO’s concept of decent work but without clarifying further about what discouraged dropouts and other dropouts mean.

Dilip Mookerjee asked whether the data could show if unemployment was voluntary or involuntary. One way of determining this would be to ask if the unemployed person had quit their last job or if they were laid off. He also asked if the data measured the duration of unemployment—short-term, chronic, or long-term.

Indira Rajaraman asked for Dr Manna’s views as a discussant for this paper on using point estimates from sample surveys, something that Dr Mahalanobis, who designed the NSS surveys, had warned against doing. She also asked the extent to which the shift from permanent NSS field survey staff to contract employees had resulted in poor training and contributed to poor data quality.

Sonalde Desai agreed that interviewer quality and training were a major issue in comparing NSS 2004–05 and 2011–12 data on the subsidiary employment status of women, which dropped significantly, and this was responsible for a large chunk of the apparent decline in female labor force participation. This drop did not affect men as much as women. The data suggests that in about a quarter of marginal farms, nobody from the household was working, which seemed unbelievable.

Pranab Bardhan raised several issues. First, the 1970 Dantwala Committee Report had recommended the use of current daily status instead of usual daily status, which is what Bardhan had used even in his 1984 book *Land, Labour, and Rural Poverty*. The problem is that the daily status data, which are better for capturing India’s unemployment and underemployment rates, are collected in terms of days rather than people. The PLFS collects current daily status but does not estimate any measures from it. The public thinks of employment and unemployment in terms of people, but both the numerator and denominator for the current daily status are in days. The better concept is actually in terms of days rather than people, but perhaps we should call it a labor utilization rate rather than an underemployment rate. He wondered if other countries, particularly Mexico, which has a large informal sector, faced the same problem, and if so, whether they used the concept of a labor utilization rate in terms of days. Related to public perceptions, Bardhan also pointed out that in comparing two rounds, the media often talks about job losses between two rounds, but that is, of course, not something that such a comparison can possibly capture.

Second, on potential underemployment, particularly for women, Bardhan pointed out that many rural women go in and out of the labor force through the year, and NSS sub-rounds should capture this seasonality. In the busy farming season, women come into the labor force, and exit in the lean season. It is important then to ask women, who report that they are not seeking work, if they would be available for work, such as tailoring and animal husbandry, if such work were available at home. It would be useful to add this question to the NSS because women's participation is extremely important for the economy. Some NSS Rounds may have added this question, but more recently, he thought, it had been dropped.

Third, Bardhan spoke about multiple subsidiary activities. He thought that the usual daily status data captured only two subsidiary activities and the current daily status captured more than two. The problem with the current daily status data was that it was captured only at the two-digit industrial code level and therefore was not easily comparable to other data. This was particularly important for subsidiary activities for women.

Fourth, Bardhan referred to the discrepancy between NSS population estimates and those from the Census. He thought that one reason for the discrepancy might be that a single household in the 2011 Census might have split into two households with a lower family size in each, and the population estimate from this household sampled say in the NSS 2017–18 would probably show a lower overall population count.

Bishwanath Goldar noted that enterprise-based employment data are not taken as seriously as household-based employment estimates and a good question to ask would be how to improve the enterprise survey-based data, especially on the quality of jobs, to make them more reliable so that they could complement each other in providing good estimates of the quality and number of jobs being created. Second, on business registers, he noted that only if such a register could be maintained accurately over time, then the employment estimates from the enterprise surveys could be used to compute an average employment per establishment and multiplied by the number of establishments in the business register to estimate overall employment in that sector.

R. Nagaraj observed that the ASI data has a separate labor schedule that is given to the Labour Bureau for processing. This schedule is probably the weakest part of ASI, but it has immensely rich data on factory workers and their earnings, which can help address the question of labor market rigidity in India. Second, there is currently a lot of emphasis on formal sector employment, and the most widely available formal sector information is the one-page data on unemployment in the *Economic Survey*. Another source

of employment data is the Labour Bureau, which processes data on all labor laws in India, which, in turn, is compiled into its annual reports, and some of it is summarized in the *Indian Labour Year Book*. However, the *Year Book* often contains obsolete and repeated data, especially on labor welfare or labor laws. This database needs to be reviewed and updated urgently.

Santanu Pramanik thought that greater use should be made of other administrative data sources in a statistical modeling framework along with survey data to come up with predictors of employment/unemployment and to produce shrinkage type estimates, which are basically weighted combinations of estimates from the NSS surveys and regression synthetic estimates from the model. This may be a better option than increasing the sample size and overburdening the survey, especially to get to disaggregated estimates at the state and district levels.

The discussion ended with the Chair, Secretary, MoSPI, Praveen Srivastava, thanking the panelists and the audience. He suggested that the Ministry of Labor and Employment should be included in discussions on employment since they are a key stakeholder. He said that a recent positive development has been that people are taking data more seriously now, and the government is also looking at data as a public good, which is why MoSPI has made data available free of charge in recent months. That is the Ministry's contribution to research. Citing an off-hand figure, he said that the Ministry earlier used to get about 200 paying users in a year and about ₹1 crore in revenue. But this has gone up to more than 2,000 users spread across the world, and if monetized, this would be a revenue of more than ₹3 crores just over the preceding two months.

He also said that a huge cost is attached to data generation in terms of manpower, resources, training, and capacity building. He noted that the Ministry is improving its training programs and revamping its systems. It has created a Division on Data Quality Assurance to improve data collection right from the capturing stage. This would help resolve a lot of issues over the next few years. While people criticize the EPFO data, they need to know that the data being captured from September 2017 onwards is uniquely identifiable to one particular person which prevents any duplication. The Ministry of Labour has undertaken a massive exercise to de-duplicate the EPFO database. He requested the research community to look at various administrative datasets available, the survey data available, and different elements in the large matrix of the time-use survey and the EC, which would be updated every three years. The Ministry also plans to integrate elements from the GSTN, the EPFO, the ESIC, and the database of the Ministry of Corporate Affairs.

The Ministry also held a session with several research organizations to discuss the UN's Sustainable Development Goals. It has come out with a framework of 300 indicators, but this covers only 80 of the 230-odd indicators assessed globally, thereby leading to a huge data gap. He finally noted the establishment of a new NSC, among other measures for strengthening institutions in the data ecosystem.

Shekhar Shah asked about the implications of the merger of the NSSO and the Central Statistics Office (CSO) and whether that raised the issue of the independence of the statistical establishment in India, which has been taken for granted for the last 70-odd years.

Praveen Srivastava noted that there was no dilution of any institution's autonomy and said that the merger of the NSSO and CSO was a logical step that should have happened a long time back. This is because there are a lot of commonalities in the work of the two institutions. The data produced by NSSO is used by CSO, and some of the results of the CSO are used to corroborate the data of NSS, and bringing them to a common entity was global best practice. The merger will bring synergies within the existing resources, including both survey data and administrative data. This would also ensure that all stakeholders work in the right direction for making data available as a public good.

The session video, the paper, and all presentations for this IPF session are hyperlinked on the IPF program available on the NCAER website by scanning this QR code or going to the following URL:



[https://www.ncaer.org/IPF2019/Agenda/Agenda\\_IPF\\_2019.pdf](https://www.ncaer.org/IPF2019/Agenda/Agenda_IPF_2019.pdf)