# The Periodic Labour Force Survey and the Estimate of the Socio-economic Inequalities

## A Critical Examination

MANOJ JATAV, DEEPIKA JAJORIA

A new system of data collection on issues pertaining to labour and employment, called the Periodic Labour Force Survey, replacing the very comprehensive and detailed surveys on the employment and unemployment situation, also known as quinquennial surveys, was introduced in India in 2017. This paper is an attempt to highlight the effects of the modified sampling methods adopted in the PLFS on data outcomes and inconsistencies. Compared to the EUS survey, the determining criteria used in the PLFS for classifying households across various socio-economic strata seem to be irrational, less comprehensive, and technically incorrect. A fundamental change in the basis of sample selection introduced with the PLFS makes it incomparable to the earlier surveys.

The authors are indebted to S K Sasikumar for providing insightful comments and suggestions in improvising the manuscript.

Manoj Jatav (*jatavjnu@gmail.com*) is an associate fellow, V V Giri National Labour Institute, Noida. Deepika Jajoria (*deepikaomatic@gmail.com*) teaches economics at Shyamlal College, University of Delhi.

Labour and employment outcomes of the first annual estimates (2017–18) under the system of the Periodic Labour Force Survey (PLFS) have been striking and contentious. Unlike the past household-level sample surveys on the employment and unemployment situation (EUS) conducted by the National Sample Survey Office (NSSO), results from the PLFS data indicate a remarkably exacerbating performance of key labour market indicators, such as the labour force participation rates (LFPRs), the workforce participation rates (WPRs), and unemployment rates. So far, four annual rounds have been conducted under the system of the PLFS from 2017–18 to 2020–21. The trends in the labour market recall the concerns of economists and demographers about the potential "unfortunate lost opportunity" as the demographic dividend of India seems to have been not utilised by the policymakers to boost the per capita income growth (Bhalla et al 2017). Broadly, it can be stated that India has witnessed a drastic increase in the unemployment rates and the share of persons of the total working-age (or economically most productive) population, that is aged 18–64, who are "Not in Employment, nor in Education, nor in Training" (NEET) between 2011–12 and 2017–18. Despite rising numbers in the labour force, there has been a substantial decline in the LFPRs and the WPRs, along with a significant surge in the unemployment and NEET rates, particularly in the young age group (18–29). A similar trend was observed between 2004–05 and 2011–12, but in the more recent period, it has been even sharper. The observed trends are starker in rural areas, where the size of the workforce (18–64 age group) dropped by 7 million from 2011–12 to 2017–18, accompanied by a dramatic increase in the unemployment and NEET rates by 3.6 and 7.7. The second annual estimates in 2018–19 also showed nearly similar levels as observed in the first annual estimates during 2017–18 (Table 1, p 36).

Such alarming changes observed in the employment statistics in India were discussed in academic, media and policy circles, particularly the credibility of data estimates obtained from the PLFS (Kapoor 2018; Mitra and Singh 2019; Saratchand 2019). As part of the "statistical reforms" that took place in India since 2014, the system of collecting socio-economic data through the NSSO's EUS surveys has undergone crucial changes as far as the methodology of collecting data and the sampling design are concerned (Jatav and Jajoria 2019; Kapoor 2018). A task force on improving employment data was constituted by the

**Table 1: Trends in Major Indicators of the Labour Market in India among the 18–64 Age Group Population**

| Sector | Survey Year | Labour Force (million) | Workforce (million) | LFPR | WPR | UR | | NEET | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Age 18–29 | Age 18–64 | Age 18–29 | Age 18–64 |
| Rural | 2004–05 | 327 | 322 | 74 | 73 | 4.0 | 1.6 | 28.1 | 24.8 |
| | 2011–12 | 328 | 323 | 65 | 64 | 4.6 | 1.5 | 32.6 | 30.7 |
| | 2017–18 | 333 | 316 | 58 | 55 | 15.9 | 5.1 | 41.3 | 38.4 |
| | 2018–19 | 340 | 324 | 59 | 56 | 15.1 | 4.7 | 41.2 | 37.5 |
| Urban | 2004–05 | 100 | 95 | 59 | 56 | 10.0 | 4.3 | 34.6 | 37.2 |
| | 2011–12 | 124 | 120 | 55 | 53 | 8.9 | 3.3 | 33.0 | 38.0 |
| | 2017–18 | 140 | 129 | 54 | 50 | 20.3 | 7.7 | 37.2 | 41.0 |
| | 2018–19 | 148 | 137 | 54 | 50 | 19.7 | 7.5 | 36.8 | 40.9 |

LFPR—number of persons in the labour force per 100 population; WPR—workers per 100 population; UR—number of unemployed persons per 100 labour force; NEET—share (%) of "Not in Employment, nor in Education, nor in Training" in total population. Labour force and workforce (in millions) are adjusted census population projected for the corresponding years.
Source: Authors' calculations using NSSO data (usual and subsidiary activities combined).

NITI Aayog, which realised the need to overcome the shortcomings of the NSSO's surveys on the EUS—most importantly, the low frequency of data and time-lag between the collection of data and availability of results. Hence, the NSSO replaced the EUS survey with the PLFS. The first annual survey under the PLFS was conducted during 2017–18. After the release of this first annual survey data, the PLFS has been criticised widely, particularly on the grounds of credibility of data and modified survey methods adopted.

This paper is an attempt to probe the effects of the modified survey methods on data outcomes. It does so by comparing the newly introduced PLFS 2017–18, with the previously conducted EUS 2011–12. The paper is divided into six parts. The first three parts contextualise the study, set out the objectives, and discuss the data, concepts and methods used. The next two, examine the consistency and comparability of the data estimates generated from the PLFS and EUS. The paper ends with the findings and suggestions.

## EUS Survey: Brief History and Significance

The NSSO's EUS survey is significant in terms of coverage and scope. It is a comprehensive survey that provides detailed statistics not only of key indicators of employment, but also other parameters which are relevant for studying socio-economic inequalities and inclusiveness, the status of decent work, various dimensions of unemployment, wages, and so on (Kannan and Raveendra 2019; Papola 2014). The initial surveys were irregular, lacking appropriate representation of samples across seasons, regions and communities. In due course, the Dantwala Committee of experts was set up in August 1968 by the erstwhile Planning Commission of India to develop concepts and definitions for a systematic survey (NSO 2019; Papola 2014). It was the beginning of a new EUS survey on a regular basis—usually with a frequency of five years (quinquennial basis) that jettisoned the inconsistencies and errors of the previous surveys.

The first quinquennial survey (27th round) was conducted between October 1972 and September 1973, covering 365 days. The survey period was divided into four quarters (of three months each) along with the allocation of an equal number of households to be surveyed in each quarter so that the survey could cover the seasonality of the EUS in both lean and abundant/peak seasons. Apart from this, robust and improved sampling techniques were introduced to assure adequate representation of different socio-economic groups and regions. The NSSO put considerable efforts in enhancing the sampling frame, sampling design, and methodology for sample selection, codification of the survey questionnaire, data collection and data processing. After 2011–12, the EUS survey was discontinued and modified into a new format called the PLFS.

Following international norms, the International Labour Organization conventions, and various Sustainable Development Goals, India's development policy has focused on two broad objectives: poverty reduction and inclusive development. Therefore, availability of regular, robust and reliable set of data on various socio-economic parameters, particularly on unorganised labour and other vulnerable sections in the population—such as Scheduled Castes, Scheduled Tribes, and women is imperative. The quinquennial surveys provided information on a range of indicators pertaining to persisting labour market inequalities and status of decent work, along with other relevant aspects such as prevailing poverty, socio-economic inequalities, gender inequalities, and educational inequalities. The NSSO, in its EUS surveys, followed robust multistage stratified random sampling techniques to capture these inequalities. However, there are several disadvantages of the PLFS system as it deviates from the purpose of capturing various dimensions of inequalities.

## Data

This paper is based on three household-level NSSO surveys—the 68th round conducted during July 2011–June 2012 and two annual surveys under the PLFS scheme conducted during July 2017–June 2018 and July 2018–June 2019. The surveys were similar to each other in terms of methods used in determining and allocating sample sizes across the states and districts in rural and urban sectors. The allocation of sample sizes at state and district levels is done by their actual proportion as reflected in the census population survey. The share of rural and urban areas in the total sample is determined in the same way (NSO 2019; NSSO 2014). The only dissimilarity observed is in the criterion adopted for the final selection of the households to be sampled. This creates a fundamental difference between the two data sets (EUS and PLFS) as it is expected to have a substantial effect on the overall data outcomes, particularly while measuring the inequalities from it.

## Selection of the Sample Households

In both the EUS and PLFS, the selection of the sample households is done in four stages. First, the village directory of the Census of India and the Urban Frame Survey (UFS) are utilised to list the first state units (FSUs) identified as villages and urban blocks respectively (NSO 2019; NSSO 2014). Second, the FSUs are further divided into hamlet group/s or sub-block/s in rural and urban areas respectively based on the size of the population in the FSU. Third, the households in each FSU and its hamlet group/s or sub-block/s (if any) are classified into three second-stage strata (SSS). Finally, the sample households are identified in the fourth stage to form the ultimate stage units (USUs). The numbers of FSUs and sample households for EUS 2011–12, PLFS 2017–18 and 2018–19 are set out in Table 2.

**Table 2: Sample Size (in Thousand) in EUS 2011–12 and PLFS 2017–18 and 2018–19**

| Survey | Total Number of FSUs | | | Households Surveyed* | | |
|---|---|---|---|---|---|---|
| | Rural | Urban | Total | Rural | Urban | Total |
| PLFS 2018–19 | 7.0 | 5.7 | 12.7 | 55.8 | 45.8 | 101.6 |
| PLFS 2017–18 | 7.0 | 5.8 | 12.8 | 56.1 | 46.0 | 102.1 |
| EUS 2011–12 | 7.5 | 5.3 | 12.7 | 59.7 | 42.0 | 101.7 |

* For PLFS, it is the number of households surveyed during the first visit.
Source: Authors' compilation from NSSO reports (Jajoria and Jatav 2020).

The categorisation of households by the SSS is the most crucial stage of the sampling procedure as it has a direct bearing on the various dimensions of inequality. While stratifying the sample, the basic rule to be followed is to ensure that the strata are essentially independent and mutually exclusive subsets of the population; also, heterogeneity among strata and homogeneity within strata guide the establishment of a stratum (UN 2008; Yansaneh 2005). The formation of the SSS before the final selection of the households helps minimise bias in the sampling procedure while ensuring appropriate representation of each socio-economic group in the population (UN 2008). In the PLFS, the educational backwardness (or forwardness) among the household members has been taken as an indicator of the formation of the SSS, which replaces the criterion adopted in the previous EUS surveys. Previously, the SSS was formed with the help of "household-listing" mentioning the level of economic affluence of the households (NSSO 2014). The criterion followed in the PLFS appears to be simplistic and does not take into account the actual economic status of the households. It could have been treated as a macro indicator of development, rather than a criterion for sample household selection.

In the previous survey system, two separate criteria were adopted for rural and urban areas to determine who is relatively affluent (Table 3). For that, a detailed exercise used to be performed by the field investigators to identify the households. For rural areas, the relatively affluent households were identified by taking into account the factors generally associated with richer households, such as ownership of the farm, non-farm, and household assets, of a large business or working in a highly remunerative profession, of a spacious *pucca* house in good condition, and so on (NSSO 2014).[1] In urban areas, households were categorised across the SSS based on the level of monthly per capita consumer expenditure (MPCE). In both the EUS and PLFS, the coding of the SSS is done in such a way that the first category of households represents the better-off households while the middle and the bottom categories broadly represent the middle-income and the poorest households respectively.

**Table 3: Criteria for SSS Formation and Sample Household Selection**

| SSS | Rural Areas | Urban Areas |
|---|---|---|
| PLFS 2017–18 and 2018–19 (education) | | |
| 1 | Two or more members of the household educated up to secondary school or more | Three or more members of the household educated up to secondary school or more |
| 2 | One member of the household educated up to secondary school or more | Two members of the household educated up to secondary school or more |
| 3 | No member of the household educated up to secondary school or more | One member of the household educated up to secondary school or more |
| 4 | * | No member of the household educated up to secondary school |
| EUS 2011–12 (economic affluence) | | |
| 1 | Relatively affluent households | Households having MPCE of top 10% of urban population |
| 2 | Of the remaining, households having principal earnings from non-agricultural activity | Households having MPCE of middle 60% of urban population |
| 3 | Other households | Households having MPCE of bottom 30% of urban population |

* Stratum does not exist for rural areas.
Source: Authors' compilation (NSO 2019; NSSO 2014).

The study reviews the potential effects of the modified sample selection method on the data outcome. It is essential that stratification (SSS) is also reflected in the data and correspond to the existing socio-economic inequalities in the population appropriately. In addition, the general indicators of development, such as the level of education and MPCE, move in a positive direction with the SSS. Further, inconsistencies in the PLFS estimates are examined by comparing it with the last quinquennial survey. This is done across the 20 vigintiles/demi-deciles (abbreviated as "v") formed by classifying the households based on the level of MPCE. The MPCE of a household is widely used as a proxy indicator of the household's income. In the quinquennial surveys, the MPCE is calculated based on the responses to detailed queries in the survey questionnaire on a household's consumer expenditure incurred on various perishable (non-durable) and non-perishable (durable) goods and services for two different reference periods: the last month and year. However, such detailed queries on MPCE are replaced in the PLFS with merely a query on "household's *usual* monthly expenditure" in the survey questionnaire (Jajoria and Jatav 2020; NSSO 2014, 2019). Therefore, the PLFS estimates are expected to show lower levels of MPCE compared to the quinquennial rounds due to the exclusion of the expenditure incurred on durable goods and services. Nevertheless, classification of the households across the 20 vigintiles can be done in both the EUS and PLFS using the existing data on the MCPE.

In the EUS, the survey period is divided into four quarters (of three months each), which ensures equal representation of all the seasons. Unlike the previous surveys, the PLFS provides data on an annual basis. In addition, a "panel data collection" scheme has also been introduced for urban areas; through this, one would be able to estimate the periodic changes in the key labour force indicators on a quarterly-cum-annual basis (NSO 2019). However, the scheme has been adopted only for urban areas. In the rotational panel design, a household will be visited four times, the first time during the first visit (or the quarter) and the other three times during the revisits (the next three quarters). This could be considered a significant value addition in the PLFS. Though the PLFS ensures a regular supply of data for both rural and urban areas, it does not compensate for the major omissions in its questionnaire, particularly the queries used by the researchers as covariates along with the labour data. These include access to land owned, cultivated and irrigated, principal economic activity and occupation of the household, detailed queries related to consumer expenditure behaviour of the households, quality of the workforce, domestic duties and nature of unemployment (Jatav and Jajoria 2019).

## Differences in the Criterion of SSS Determination

In the previous surveys, the NSSO classified the households across three SSS categories, in both rural and urban areas. The PLFS, by introducing the education criterion instead of using the one based on the MPCE as in the previous surveys, has four categories of SSS in the urban areas. Thus, for urban areas, there would be no possibility of making comparisons between the two data sets. We look at the distribution of the estimated number of households (based on projections made on the sample data)[2] in the PLFS 2017–18 and 2018–19 by the SSS and compares it with two previously conducted quinquennial rounds: EUS 2004–05 (61st round) and 2011–12 (68th round). The purpose of showing the distribution in EUS 2004–05 and 2011–12 is to examine the existence of any inconsistency across the EUS surveys. It also seeks to see whether the distribution in the PLFS corresponds to that in the EUS surveys. Based on the results, it can be stated that the distribution in the PLFS appears to be unrealistic and incomparable with the EUS surveys. It is important to critically assess education as a criterion for identifying the sample households across the strata, as it does not appear to be a robust one for several reasons, which are discussed next.

A comparison of the EUS rounds (2004–05 and 2011–12) reveals that the change in the estimated distribution of the households by the SSS was consistent as there was a gradual increase in the percentage shares of the middle (broadly representing the middle-income households) and upper (relatively affluent households) strata, accompanied by a decrease in the share of the bottom SSS (broadly representing the relatively poorer households), both in rural and urban areas (Table 4). A shift from the bottom stratum to the middle and upper strata indicates consistently rising levels of households' income along with the process of economic growth. However, in the PLFS estimates, an unprecedented shift from the bottom stratum to the upper stratum is observed (that is from 6.5% during 2011–12 to 26.1% and 27.1% during 2017–18 and 2018–19, respectively in rural areas), accompanied by a substantial decline in the share of the middle stratum. The sudden increase in the share of relatively affluent households in the estimation looks unrealistic, as it indicates that the income of rural households increased rapidly between 2011–12 and 2018–19. On the contrary, there has been an economic slowdown during the same period accompanied by an unprecedented job-loss scenario in rural areas. From 2005 to 2009, the country witnessed a mean annual growth of 8.2% in the real GDP, which decreased to 7.2% from 2010 to 2014 and further to 6.7% from 2015 to 2019.[3]

Table 5 attempts to uncover the inconsistencies by comparing the levels of mean MPCE of the households in rural and urban areas. To compare mean MPCE levels in the EUS and PLFS surveys, the data is normalised at constant prices for 2011–12. However, the results should be cautiously interpreted due to the difference in the criteria of recording MPCE, as mentioned previously. The results show an increase in the overall MPCE levels among rural households along with reduced levels of inequality measured with the help of Gini's index of inequality. The gap in the mean MPCE between the upper and middle strata seems to be substantially reduced from ₹682 during 2011–12 to ₹233 during 2017–18 and further reduced to ₹191 during 2018–19. In addition, a reduction in the mean MCPE by ₹67 was noticed between 2011–12 and 2017–18 in the upper

**Table 4: Percentage Distribution of Samples and Estimation by SSS in Various Surveys Conducted by the NSSO** (%)

| Survey | SSS Number | Rural | | Urban | |
|---|---|---|---|---|---|
| | | Estimated Households | Sampled Households | Estimated Households | Sampled Households |
| PLFS 2018–19 | 1 | 27.1 | 24.9 | 22.5 | 25.2 |
| | 2 | 24.3 | 49.9 | 24.0 | 24.9 |
| | 3 | 48.6 | 25.1 | 23.2 | 25.5 |
| | 4 | NA | NA | 30.2 | 24.4 |
| | All | 100 | 100 | 100 | 100 |
| PLFS 2017–18 | 1 | 26.1 | 24.9 | 22.1 | 25.2 |
| | 2 | 23.7 | 49.9 | 23.7 | 24.9 |
| | 3 | 50.3 | 25.1 | 23.2 | 25.5 |
| | 4 | NA | NA | 31.1 | 24.4 |
| | All | 100 | 100 | 100 | 100 |
| EUS 2011–12 | 1 | 6.5 | 24.7 | 22.9 | 28.9 |
| | 2 | 36.5 | 49.6 | 65.9 | 52.0 |
| | 3 | 57.1 | 25.7 | 11.3 | 19.1 |
| | All | 100 | 100 | 100 | 100 |
| EUS 2004–05 | 1 | 6.2 | 20.0 | 19.4 | 23.0 |
| | 2 | 31.7 | 39.0 | 62.7 | 42.2 |
| | 3 | 62.1 | 41.1 | 17.8 | 34.8 |
| | All | 100 | 100 | 100 | 100 |

NA—not applicable (SSS does not exist for rural sector).
Source: Authors' calculations using NSSO data.

**Table 5: Estimated Inequality in MPCE across SSS**

| Survey | SSS | Mean MPCE* | | Gini Coefficient^ | |
|---|---|---|---|---|---|
| | | Rural | Urban | Rural | Urban |
| PLFS 2018–19 | 1 | 2,052 | 3,969 | 0.29 | 0.33 |
| | 2 | 1,861 | 4,125 | 0.29 | 0.35 |
| | 3 | 1,568 | 3,951 | 0.26 | 0.40 |
| | 4 | NA | 2,670 | NA | 0.34 |
| | Total | 1,770 | 3,610 | 0.29 | 0.37 |
| PLFS 2017–18 | 1 | 1,969 | 3,874 | 0.32 | 0.35 |
| | 2 | 1,737 | 3,927 | 0.30 | 0.37 |
| | 3 | 1,478 | 3,623 | 0.28 | 0.40 |
| | 4 | NA | 2,443 | NA | 0.35 |
| | Total | 1,667 | 3,384 | 0.30 | 0.38 |
| EUS 2011–12 | 1 | 2,036 | 5,169 | 0.35 | 0.32 |
| | 2 | 1,354 | 2,138 | 0.30 | 0.27 |
| | 3 | 1,208 | 1,032 | 0.28 | 0.22 |
| | Total | 1,315 | 2,706 | 0.30 | 0.39 |

* INR at 2011–12 constant prices; NA—not applicable (SSS does not exist for the rural sector);
^ ranges from 0 (complete equality) to 1 (complete inequality).
Source: Authors' calculations using NSSO data.

stratum; however, it again increased by ₹83 in only one year during 2017–18 and 2018–19. These results are subject to scrutiny given the prevailing levels of inequality and rising wealth gap observed over this period in rural areas of India. The other possibility is that there is a statistical error involved in the sample selection procedure, wherein the households representing characteristics of the middle stratum have been included in the upper stratum. Therefore, it is likely that the reduced differences in the mean incomes across various strata have occurred due to the overlap in the middle and upper strata, with more or less similar household characteristics in the total sample. There is a high chance that the units included in the middle and the upper strata have similar characteristics, creating an overlap in the characteristics across the strata. Resultantly, there is higher heterogeneity within the middle and upper strata and higher homogeneity across various strata. As previously mentioned, the homogeneity within and heterogeneity across various strata is essential for a sample to provide a good representation of the population; the abovementioned statistical error makes the PLFS data a poor representative of the population. Sample household determination based on the education criterion is too mechanical to capture the real economic status of the households.

The causal link between education and economic well-being runs both ways. A higher level of economic well-being makes access to education easy and a higher level of education serves as one of the various means to attain a higher standard of living. While choosing education as an indicator of economic well-being, we are considering economic well-being as an end and education as a means to this end. A higher level of education makes the individual capable to access better jobs and higher levels of earning, thereby improving the material well-being of the individual. There is a chain of links between education and standard of living—for a higher level of education to serve as a means to attain a higher level of material well-being. And, the availability of employment opportunities is an important link between the two. The weakening access to jobs due to lack of employment opportunities also weakens the link between the level of education and standard of living. Although there is a positive correlation between the level of education and standard of living, the correlation is not high enough to make the former a proxy indicator of the latter. Further, due to its mechanical nature, there is no probability that the economic status of the households during the pre-listing survey is verified by the field investigator.

While the PLFS estimates for urban areas are not comparable with the results of the previous survey, the mean MPCE of the upper stratum is reduced unprecedentedly from ₹5,169 during 2011–12 to ₹3,874 during 2017–18 (Table 5). In addition, the difference of the mean MPCE across the strata (SSS one to SSS four) appears to be very small, particularly in the first two strata. This also indicates an overlap between the middle and upper strata, leading to biased estimation from the data, with over-representation of middle-income households (and under-representation of the real upper stratum) in the sampling procedure of the PLFS originating from the ambiguous educational criterion.

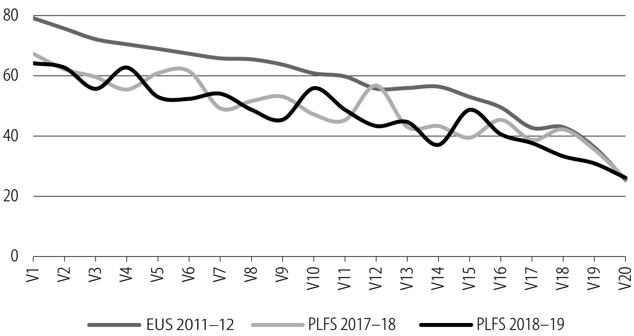
## Is Education a Better Criterion for SSS Formation?

The determination of sample households across strata needs to be based on indicators that are directly indicative of a household's access to various assets of livelihoods. The NSSO had used different criteria for rural and urban areas, that is economic affluence and the level of MPCE respectively. Both the criteria are based on the same principle. To categorise the households by the SSS in rural areas, access to various assets of the households, both household and business assets, is determined. Whereas for categorising the households in urban areas, the MPCE, which is a proxy for a household's income, is used. The MPCE is also, more or less, directly indicative of a household's economic condition. On the other hand, in the PLFS, the NSSO has universalised the basis of household selection, both in rural and urban areas, by replacing the previous criteria with one based on educational attainments among the household members. The educational criterion is defined as "the number of household members educated up to secondary school or more." This criterion has a few fundamental issues that can lead to serious specific errors during household selection across strata. First, it does not measure a household's economic status directly; rather, it can be considered an indirect indicator affirming the economic progress of the household. Further, better education is only a means to the achievement of a better standard of living; it is not an end (Jatav and Jajoria 2019).

Second, it might exclude those households from the upper stratum where less than two or three members fulfil the criteria. For example, such a problem can arise in the case of small households, particularly in urban areas where the mean size of the household is comparatively smaller than that in rural areas, and small households with children who do not fulfil (or are not eligible to fulfil) the household selection criterion due to age constraints. Thus, there is a higher probability of such households not being counted in the upper stratum due to the defined threshold of the number of household members and education level, even though they otherwise possess the required characteristics.

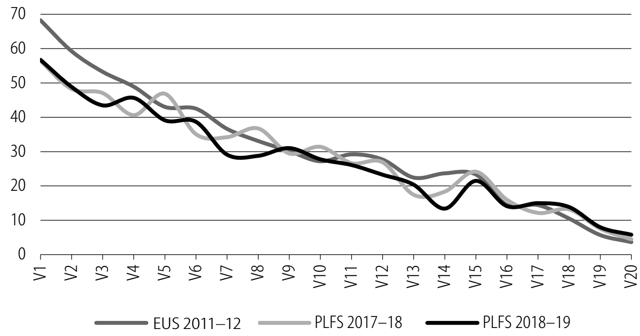
Third, due to a higher prevalence of fertility rates and a resultant bigger household size among the lower socio-economic stratum, there is a possibility that these households would be counted in the middle or upper stratum, even though they have a disadvantageous position in the society. Due to an increase in the mean years of schooling over time (UNDP 2016), the probability of being counted in the higher socio-economic stratum has increased among relatively poor households. Further, having attained education up to Class 10 is not an adequate indicator of the achievement of a better standard of living when there are high rates of underemployment and unemployment. Therefore, there is a chance that the new criterion might lead to miscounting of the households belonging to the lower-most and upper-most strata (Jatav and Jajoria 2019).

**Figure 1: MPCE Class-wise Estimated Percentage of Rural Households Having No Member with Secondary (or More) Level of School Education**



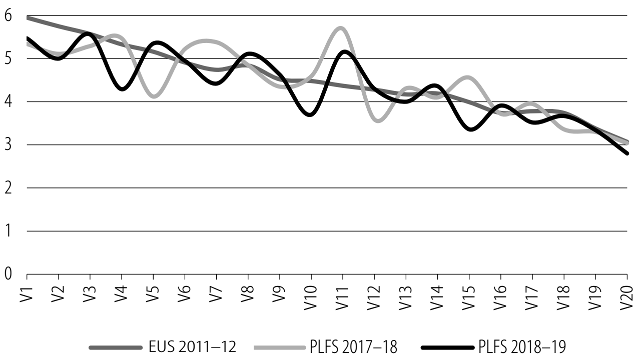

Source: Authors' calculations.

**Figure 2: MPCE Class-wise Estimated Percentage of Urban Households Having No Member with Secondary (or More) Level of Schooling**



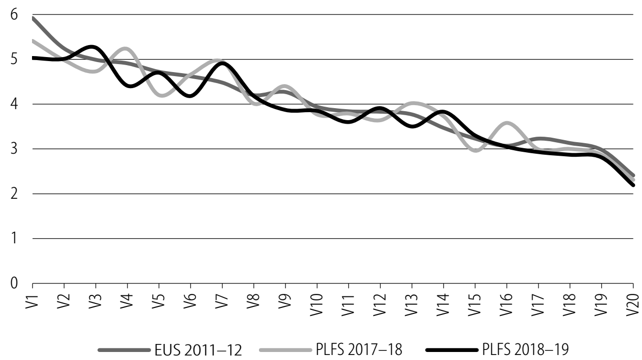

Source: Authors' calculations.

**Figure 3: MPCE Class-wise Estimated Mean Household Size in Rural India**




Source: Authors' calculations.

**Figure 4: MPCE Class-wise Estimated Mean Household Size in Urban India**




Source: Authors' calculations.

Similarly, any household can be counted in any stratum due to the arbitrariness of the selection criteria. This is shown in Table 6. The estimated distribution of households' PLFS data (for 2017–18 and 2018–19) suggests that the average number of household members with education up to secondary school or more increases with a rise in the level of MPCE—though there are fluctuations observed across the levels of MPCE, which indicates existing inconsistencies in the data. However, there is a greater chance that most of the households do not have the required threshold number of members with the defined threshold level of education across levels of MPCE, more evidently in the rural areas (see the distribution of mode; Table 6). In other words, a higher level of MCPE does not consistently correspond to better educational attainment among the household members and vice versa. In addition, despite these specific errors, the percentage share of households having no member fulfilling the required threshold of education declines with an increase in the level of MPCE in both PLFS (2017–18 and 2018–19) and EUS 2011–12. However, for both rural and urban areas, such a pattern does not show a consistent decline across levels of MPCE in the PLFS data compared to that in the EUS data (Figures 1 and 2). In both rural and urban areas, there is a linear pattern observed across the levels of MPCE in the EUS data, while the PLFS data shows non-linearity in this pattern.

**Table 6: Number of Persons in the Household with Education up to Secondary School or More, PLFS 2017–18 and 2018–19**

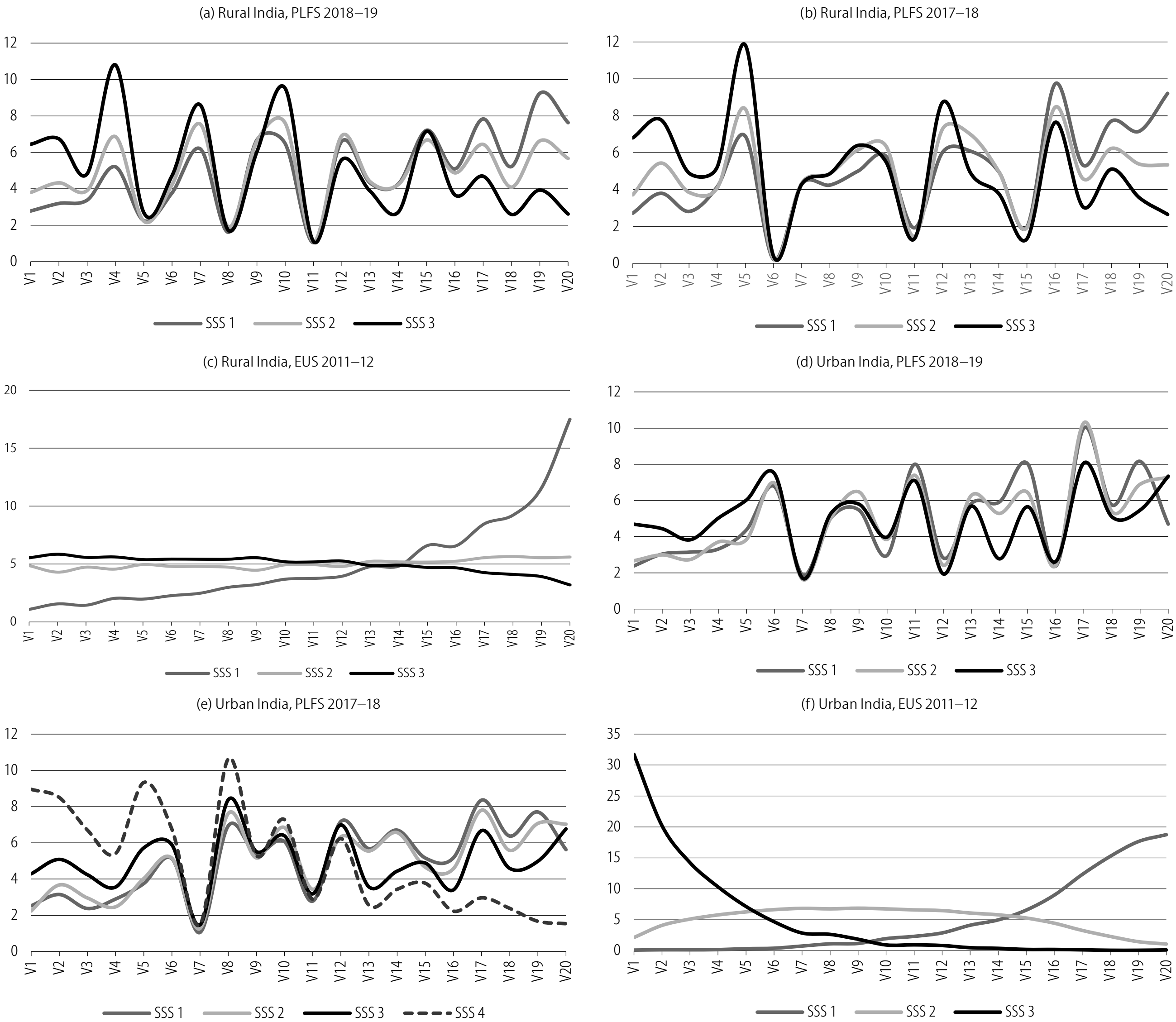
| MPCE | Rural | | | | Urban | | | |
|---|---|---|---|---|---|---|---|---|
| | $\bar{x}$ | | Mo | | $\bar{x}$ | | Mo | |
| | 2017–18 | 2018–19 | 2017–18 | 2018–19 | 2017–18 | 2018–19 | 2017–18 | 2018–19 |
| V1 | 0.6 | 0.6 | 0 | 0 | 0.9 | 0.9 | 0 | 0 |
| V2 | 0.6 | 0.6 | 0 | 0 | 1.0 | 1.1 | 0 | 0 |
| V3 | 0.7 | 0.8 | 0 | 0 | 1.0 | 1.2 | 0 | 0 |
| V4 | 0.8 | 0.6 | 0 | 0 | 1.3 | 1.1 | 0 | 0 |
| V5 | 0.7 | 0.9 | 0 | 0 | 1.1 | 1.3 | 0 | 0 |
| V6 | 0.7 | 0.8 | 0 | 0 | 1.4 | 1.3 | 0 | 0 |
| V7 | 0.9 | 0.8 | 0 | 0 | 1.4 | 1.7 | 0 | 0 |
| V8 | 0.9 | 0.9 | 0 | 0 | 1.3 | 1.6 | 0 | 0 |
| V9 | 0.8 | 1.0 | 0 | 0 | 1.6 | 1.5 | 0 | 0 |
| V10 | 0.9 | 0.8 | 0 | 0 | 1.5 | 1.5 | 0 | 0 |
| V11 | 1.1 | 1.0 | 0 | 0 | 1.6 | 1.7 | 2 | 0 |
| V12 | 0.7 | 1.0 | 0 | 0 | 1.6 | 1.9 | 0 | 2 |
| V13 | 1.0 | 1.0 | 0 | 0 | 2.0 | 1.8 | 2 | 2 |
| V14 | 1.0 | 1.2 | 0 | 0 | 1.9 | 2.1 | 2 | 2 |
| V15 | 1.2 | 0.9 | 0 | 0 | 1.7 | 1.9 | 2 | 2 |
| V16 | 1.0 | 1.1 | 0 | 0 | 2.0 | 1.8 | 2 | 1 |
| V17 | 1.2 | 1.2 | 0 | 0 | 2.0 | 1.9 | 2 | 2 |
| V18 | 1.1 | 1.3 | 0 | 0 | 2.0 | 2.0 | 2 | 2 |
| V19 | 1.2 | 1.4 | 0 | 0 | 2.1 | 2.1 | 2 | 2 |
| V20 | 1.5 | 1.4 | 2 | 1 | 1.9 | 1.8 | 2 | 2 |
| Total | 0.9 | 0.9 | 0 | 0 | 1.6 | 1.6 | 0 | 0 |

$\bar{x}$—Mean; Mo—Mode.
Source: Authors' calculations.

## Consistencies in Data Sets

We now examine the consistency and comparability of the data estimates generated from the PLFS and EUS. Results reveal that several inconsistencies exist in the PLFS data, making it incomparable with the EUS data. It is observed in the distribution of households by size in each MPCE class. Conventionally, the household size tends to decline with rising levels of per capita income in a linear and decipherable way (Lanjouw and Ravallion 1995).

**Figure 5: Percentage Distribution of Estimated Households by MPCE Class across Each SSS**

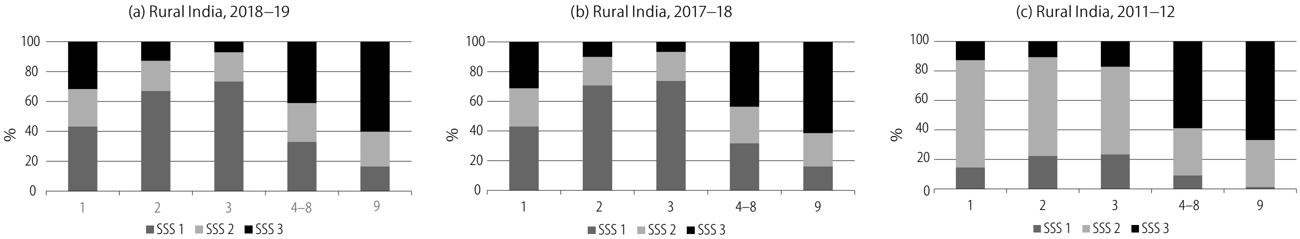



Estimates obtained from the EUS illustrate the same, both in rural and urban areas. However, it cannot be propounded by the PLFS data due to fluctuating (non-linear) mean household size across the levels of MPCE, both in rural and urban areas (Figures 3 and 4, p 40).

In a consistent data set, the distribution of households by levels of MCPE corresponds to the appropriate share of each MPCE class across the three strata. From both the PLFS and EUS data, the percentage share of each MPCE group is calculated across the three strata separately in both rural and urban areas (Figures 5a–d). It is expected that households from the bottom socio-economic stratum would have a greater concentration in the lower MPCE classes and vice versa. Similarly, the upper stratum is expected to have a greater concentration in the higher MPCE classes and vice versa; and the middle stratum would show a greater concentration in the middle MPCE levels. Thus, in the case of a consistent data set, a decipherable linear pattern is present when looking at the share of each MPCE group across each stratum. As anticipated, both in rural and urban areas, the PLFS data estimates reveal an inconsistent (non-linear) pattern of distribution of households (percentage share) by levels of MPCE across all the strata (Figures 5a, b, d, and e). The EUS data estimates, on the other hand, were found to be consistent across all the strata in rural and urban areas (Figures 5c and f).

Despite a consistent pattern across the three strata in EUS data, the middle stratum in rural areas seems to have a uniform distribution of the levels of MPCE (Figure 5c). This points to the potential caveats in the quinquennial rounds of the NSSO while defining the middle stratum in rural areas. As discussed earlier, households in the middle stratum were defined as those with principal earnings from non-agricultural activities, excluding the relatively affluent ones. However, the potential of the non-farm sector to generate enough income is debatable in the context of rural India. In the quinquennial surveys, the NSSO had presumed that only those rural households primarily dependent on non-agricultural activities would have been able to satisfy the characteristics of the

**Figures 6: Percentage Distribution of Estimated Workers (UPSS) by SSS within Each Occupational-cum-Skill Category**




Source: Authors' calculations.

middle stratum. It excluded agricultural households from the middle stratum and, thus, generated biased data estimates. It excluded those agricultural households with less than 7 hectares of cultivable land or 3.5 hectares of irrigated land from the middle stratum (NSSO 2014). Further, the "non-agricultural income" criterion is not a good fit to decide the middle rural stratum since non-agricultural income varies widely across different geographies and community groups in rural India.

In PLFS 2017–18 and 2018–19, in comparison to EUS 2011–12, another type of inconsistency is observed while examining the percentage share of workers by each stratum within various occupational-cum-skill categories in the rural areas (Figures 6a, b, and c). In the occupational hierarchy, workers belonging to the bottom stratum tend to be highly concentrated in the lower occupations, whereas the middle and upper strata show a relatively greater presence in the middle or higher occupations. To probe it, the broad nine occupations given in the National Classification of Occupations, 2004 have been categorised into five major groups based on the type of work (manual/non-manual) and required skill levels. In Figures 6a, b, and c, the uppermost occupational group in the hierarchy of occupations is O1, comprising legislators, senior officials and managers, whereas the elementary occupations (O9) are categorised at the bottom of the hierarchy. O2 and O3 refer to professionals and associate professionals with skill levels IV and III respectively. O4–O8 refer to service workers, shop and market sales people, skilled agricultural and fishery workers, workers in craft and related trades, and plant and machine operators and assemblers with skill level II. To define the skill levels, the years spent in formal education are taken into account. More than 15 years, 14–15 years, and 11–13 years of formal education are required for skill levels IV, III, and II respectively. Whereas up to 10 years of formal education or informal skills are required for elementary occupations.

Comparative estimates indicate a delusive increase in the percentage share of the upper socio-economic stratum along with an unprecedented decline in the percentage share of the middle stratum in all occupational-cum-skill categories from 2011–12 to 2017–18 (Figures 6a, b, and c). The maximum increase in percentage share of the upper stratum is observed in O3 (by 50.4%) followed by O2 (48.2%), O1 (28.6%), O4–O8 (22.6%), and O9 (14.8%). The middle stratum, on the other hand, experienced a striking change in the share, particularly in the upper three occupational-cum-skill categories (by -47% in O2, -39.9% in O3, -9.4% in O9, and -7% in O4–O8). Similarly, the share of the bottom stratum also declined across categories (by -15.2% in O4–O8, -10.5% in O3, -5.4% in O9, and -0.6% in O2), except in the uppermost occupation, O1, where an unprecedented addition of 18.5% is observed. Estimates from PLFS 2018–19 also correspond to the first PLFS estimates (2017–18) nearly in a similar way.

Among others, two completely unrealistic developments have been noticed between 2011–12 and 2017–18, which cannot be explained by any means or process: first, a substantial increase in the percentage share of the upper stratum in the total usual employment generated in the lowest category of occupation; second, a substantial increase in the percentage share of the bottom stratum in total employment generated in the highest category of occupation. In rural India, the occupational distribution of the workers has a strong association with the caste system, in which the historically deprived and marginalised castes are more likely to be concentrated in the lower rungs of the occupational hierarchy. It is strongly believed that overestimation of the middle stratum accompanied by underestimation of the real upper stratum has resulted in an inconsistent pattern of the distribution of workers in each occupational-cum-skill category across the three strata. In other words, the irrational criterion adopted for household-level sampling (threshold number of persons with secondary level of schooling) has resulted in biased estimates across the strata. As a result, the upper stratum has characteristics of the middle stratum.

## The Way Forward

The reliability of a nationwide sample survey lies in its representativeness: it should appropriately reflect the characteristics of different groups in the population without bias in the selection operation (UN 2008). The PLFS has not been able to fulfil this expectation. As compared to the PLFS, estimates from the EUS seem to be more reliable and comprehensive in terms of providing detailed socio-economic data covering several important aspects of inequality. The PLFS and EUS follow different criteria for selecting the sample households across the socio-economic categories (strata). The changed criterion in the PLFS is supposed to have a greater impact on data outcomes, particularly while examining socio-economic inequalities. The PLFS generates a biased and inconsistent data estimates, along with over-representation of the middle socio-economic stratum and a resultant underestimation of the upper socio-economic stratum. Thus, the stratification in the PLFS violates the basic rule of unbiased sampling across strata and adequate representation of the categories of households in consideration. The criterion of educational attainment among the household members seems to be irrational, less comprehensive and technically incorrect; it is likely to produce biased data estimates. Hence any comparison between these two data sets must be avoided.

We recommend urgent rectification of the sampling errors from the PLFS for effective and inclusive socio-economic planning in India. It can be done in two ways: (i) one approach is to revive the old structure of sampling used in the EUS surveys by rectifying the previously existing technical errors in the household selection procedure. For instance, the selection of the households for the middle socio-economic stratum needs to be revised appropriately. The middle stratum has to cover the agricultural households as well. (ii) The second way is to thoroughly review the sampling techniques adopted in the PLFS. To develop new sampling criteria for determining the socio-economic categories in the population, it is imperative to revive the detailed queries which were made in the EUS questionnaire,[4] and also develop a new comprehensive method that essentially allows the data outcomes to reflect the socio-economic inequalities that exist in reality. Rather than using a single indicator of threshold education, it must be used in an integrated way, with caste and class identities. For that purpose, levels of MPCE and social groups could be used as proxy indicators.

## NOTES

1 "Ownership of motor car/jeep/tractor/combine-harvester/truck/bus/etc; consumer durables like refrigerator/washing machine, etc; ownership of large business/highly remunerative profession high salaried income, etc; ownership of spacious pucca house in good condition; ownership of 7 hectares or more of cultivable land; ownership of 3.5 hectares or more of irrigated land; ownership of a good number of cattle, buffaloes and camels (10 or more in number)" (NSSO 2014: B-4).

2 NSSO selected a lower number of representative sample households from the relatively poorer stratum (that is the bottom stratum) in comparison to the middle-income (middle SSS) and relatively affluent households (upper stratum) for the purpose of maintaining the diversity of the population that exist in the latter two strata in term of socio-economic characteristics. For a similar reason, the share of the urban sector in total sample size is greater than that of rural.

3 https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/IND (viewed on 25 April 2020).

4 Particularly, the detailed queries on a household's consumption expenditure. "An earlier study on the Indian National Sample Survey seems to indicate that, for certain food items, a one-month reference period produces less bias than a one-week reference period" (Mahalanobis and Sen 1954; Pettersson 2005).

## REFERENCES


Bhalla, S, A Kumar and M Jatav (2017): "India's Demographic Divided: Theoretical Antecedents and Empirical Evidence," *Labour and Development: Essays in Honour of Prof T S Papola,* Rajendra P Mamgain and Preet Rustagi, K P Kannan (eds), New Delhi: Academic Foundation.

Jajoria, D and M Jatav (2020): "Is Periodic Labour Force Survey, 2017–18 Comparable with Employment–Unemployment Survey, 2011–12," *Economic & Political Weekly*, Vol 55, No 3, pp 12–16.

Jatav, M and D Jajoria (2019): "Household-level Surveys on Labour Force, Employment and Unemployment in India: A Comparative Assessment," *Labour & Development*, Vol 26, No 2, pp 29–47.

Kannan, K P and G Raveendran (2019): "From Jobless to Job-loss Growth: Gainers and Losers during 2012–18," *Economic & Political Weekly*, Vol 54, No 44, pp 38–44.

Kapoor, R (2018): "Rethinking India's Employment Data Architecture," *Economic & Political Weekly*, Vol 53, No 40, pp 14–16.

Lanjouw, P and M Ravallion (1995): "Poverty and Household Size," *Economic Journal*, Vol 105, No 433, pp 1415–34.

Mitra, A and J Singh (2019): "Rising Unemployment in India: A Statewise Analysis from 1993–94 to 2017–18," *Economic & Political Weekly*, Vol 54, No 50, pp 12–16.

NSO (2019): "Annual Report: Periodic Labour Force Survey (July 2017–June 2018)," National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

NSSO (2014): "Employment and Unemployment Situation in India," Report No 554(68/10/1), National Sample Survey Office, Government of India, New Delhi.

Papola, T S (2014): "An Assessment of the Labour Statistics System in India," International Labour Organization Country Office, New Delhi.

Pettersson, Hans (2005): "Survey Design and Sample Design in Household Budget Surveys," *Household Sample Surveys in Developing and Transition Countries,* Studies in Methods: F (96), Statistical Division, Department of Economics and Social Affairs, United Nations, New York.

Saratchand, C (2019): "Some Macro-theoretical Foundations of Jobless Growth," *Economic & Political Weekly*, Vol 54, No 13, pp 17–20.

UNDP (2016): "Human Development Report 2016," United Nations Development Programme, United Nations, New York.

United Nations (2008): "Designing Household Survey Samples: Practical Guidelines," Studies in Methods: F (98), Statistical Division, Department of Economics and Social Affairs, New York.

Yansaneh, I S (2005): "Overview of Sample Design Issues for Household Surveys in Developing and Transition Countries," *Household Sample Surveys in Developing and Transition Countries,* Studies in Methods: F (96), Statistical Division, Department of Economics and Social Affairs, United Nations, New York.