Issue. No. PDOS-57-XXXX Sept. 2024 (1-2)
ISSN 2249-197X

Data for Development

सर्वेक्षण
SARVEKSHANA

117th Issue
September, 2024

Journal of
National Statistics Office

Government of India
Ministry of Statistics and Programme Implementation
National Statistics Office
New Delhi
www.mospi.gov.in

SARVEKSHANA

117^{th} Issue

Journal of
National Statistics Office

सत्यमेव जयते

Government of India
Ministry of Statistics and Programme Implementation
National Statistics Office
New Delhi

Journal of National Statistics Office (NSO), Ministry of Statistics and Programme Implementation (MoSPI)

Frequency and Subscription
'Sarvekshana' is generally published twice a year and is made available on the website of the Ministry for free download. The subscription rate for hard copy is ₹300 per copy. For subscription mail to: Controller of Publications, Department of Publication, Civil Lines, Delhi - 110054. Ph. 23819689, 23817823

Submission of papers/articles for Sarvekshana
'Sarvekshana' is aimed at encouraging research and analysis of NSS, Census, ASI data and other official statistics to bring about a deeper understanding of statistical data. For details about submission of papers/articles, refer to back of cover page.

Suggestions for improvement of the Journal may be addressed to:
The Managing Editor, Sarvekshana,
Coordination & Quality Control Division
National Statistics Office, MoSPI
Sankhyiki Bhawan, Maharshi Valmiki Marg, Delhi – 110032

Foreword

Bringing out *Sarvekshana* has always been an enlightening endeavor. The first issue of *Sarvekshana* was released in July, 1977 and the latest released issue is combined 115th & 116th issue. The present 117th issue comes with three papers on the subjects of (i) An Alternative Imputation Technique for Treatment of 'Temporarily Missing' Food Item Prices for Compilation of Consumer Price Index (CPI), (ii) District Level Estimates of Agricultural Household Income in India Using Small Area Models and (iii) Females in the Job Market: An Understanding of Its Socio-Economic Correlates. In addition, the highlights of the recent survey report of Periodic Labour Force Survey namely, 'Annual Report of Periodic Labour Force Survey (PLFS), July 2023 – June 2024' have been included in the 117th issue.

Referees have been kind in examining the papers in detail and offering their suggestions in a short span of time. So have been the Members of the Editorial Advisory Board of *Sarvekshana*. I offer my sincere gratitude to all of them and solicit continued support for the Journal. Authors of the papers have also been very cooperative in accepting the suggestions for repetitive revisions of the papers. On behalf of the Editorial Advisory Board and on my own behalf, I congratulate them for their work which we hope would be useful. Officers of Coordination & Quality Control Division of National Statistics Office have been meticulous at various stages of publication of this issue and their hard work deserves appreciation.

The *Sarvekshana* is a known Journal among researchers, academicians and policy makers. I welcome students, researchers, Government officials and all those working on data based on sample surveys and censuses to contribute papers for this Journal.

New Delhi — Chairman

December, 2024 — Editorial Advisory Board

Contents

PART-I

TECHNICAL PAPERS

An alternative imputation technique for treatment of ‘Temporarily Missing’ food item prices for compilation of Consumer Price Index (CPI)

Kasturi Rajeswari[1]

Abstract

The National Statistical Offices (NSOs) release the Consumer Price Index (CPI) every month with a lag of about two weeks. For compilation of the same, price data are collected by NSO price collectors, from specific centres, identified markets and outlets, spread across the country, item-wise for given specifications. For reasons like, the shop being temporarily closed or specific item not available temporarily, etc., some item prices are not reported, and NSOs often face this problem of incomplete data. Such missing price data are classified as ‘temporarily (non-seasonal) missing products’ and ‘Seasonal missing products’. In this paper, multiple imputation as an alternative imputation technique for treatment of ‘Temporarily Missing’ food item prices for compilation of CPI is attempted. The performance of the conventional imputation method, adopted by most NSOs, for imputation of missing prices for compilation of the index, is assessed and compared to the Multiple Imputation method that is based on MICE algorithm. The performance will be demonstrated and compared for the two methods in compiling the ‘Food-Items’ component of Consumer Price Index, for incomplete data with MCAR and MAR missing mechanism, with varying percent of missing from 5, 10 and 15, based on the Repeated-Measures ANOVA (RM-ANOVA) analysis.

Keywords: Consumer Price Index, Inflation, Statistical Simulation, Missing data, Multiple Imputation

JEL Codes: C43, E31, D12, P22, C15

Receipt of the Final Version of paper from Author: September 2024

Acceptance: October 2024

[1]Author is PhD, an Indian Statistical Service Officer and currently working as DDG, Statistics, National Tuberculosis Institute, Bangalore

1. Introduction

Consumer Price Indices (CPIs) measure changes over time in the general level of prices of goods and services that households acquire (use or pay for) for the purpose of consumption (Graf B, 2020). CPI is a macroeconomic indicator, universally in use, as a measure of inflation, and as a tool for, inflation targeting, monitoring price stability, and as deflators in the national accounts, by governments and central banks. The indicator is often included in legislations as well in service contracts as the appropriate measure for accounting of inflation and adjusting bill payments related to wages, rents, interest, social security, pensions, etc. Therefore, CPI has a substantial and wide-ranging financial implications for governments, businesses, as well as households.

1.1 Compilation of CPI

For measuring the aggregate price changes, a sample variety of items are selected from sample outlets that represent the goods and services consumed by the households. The usual method of calculation is to take an average of the period-to-period price changes for the different products, and weights being the average amounts that households spend on them. CPIs are official statistics that are usually produced by NSOs, ministries of labour, or central banks. They are published as quickly as possible, generally within two-three weeks after the reference period.

The prices are collected from shops or other retail outlets. The price collectors are given detailed item descriptions defining each item variety, its specifications and location. The detailed specifications are included on the price collection template each period and serve as a prompt to help ensure that the same varieties are being priced. As any lack of clarity in the specifications may lead to errors, detailed checklists of variety descriptions are used. In addition, adequate attention is paid to include all details for identification of the variety and outlet on a subsequent visit. For example, along with details of the outlet location, all pertinent price-determining characteristics are included, to ensure identification of changes in quality have occurred.

For compilation of the CPI, generally, the methodology adopted is illustrated here:

(a) Firstly, the price indices are computed for elementary aggregates, known as price relatives, (ratios of current prices over base prices each item-wise). The elementary aggregate can include stratification by item variety, region and by shop type or market.

(b) The elementary item indices are aggregated using Geometric Mean (GM) of the Price Relatives of Current Prices with respect to Base Prices of different markets in consonance with the international practice.

(c) These elementary indices are the lowest level of aggregation where prices are combined into price indices. Explicit expenditure weights are available at this level of aggregation. Then using consumption expenditure as weights, associated with each level, the elementary price indices are averaged (aggregated) to obtain higher level indices. This is done here with the Laspeyres index formula.

The Laspeyres index formula in standard notations:

$$P_L = \frac{\sum_{i=1}^{n} p_i^t q_i^0}{\sum_{i=1}^{n} p_i^0 q_i^0} \quad \ldots\ldots\ldots\ldots\ldots(1)$$

Where p_i^t *=price for ith item at month t (current month)*

p_i^0 *= price for ith item at month 0, the price reference period or base period*

q_i^0 *= quantity for ith item purchased, during the reference period or base period*

n = number of items

Above formula derived equivalently as:

$$P_L = \sum_{i=1}^{n} \frac{p_i^t}{p_i^0} w_i \ldots\ldots\ldots\ldots\ldots(2)$$

Where, $w_i = \frac{p_i^0 q_i^0}{\sum_{i=1}^{n} p_i^0 q_i^0}$ *is the share of the actual expenditure on item i in period 0, also called weight of ith item*

2. Imputation in compilation of CPI

A CPI must reflect the true change in the cost of buying a fixed basket of goods and services of constant quality. The compilation of CPI often faces the problem of incomplete data. For instance, products may go temporarily missing as a result of supply shortages, due to the seller underestimating the demand or due to strikes by manufacturing or transportation workers, or issues with the supply of imported goods, etc. In these cases, the price collector, although not able to observe a price in the current period, may have obtained information (for example, from the shopkeeper) to suggest that the same variety will become available again at some-time, perhaps unknown, time in the future (Graf B, 2020).

If it is believed that a missing product will be available again in a reasonable time, then the price is termed as ‘temporarily (non-seasonal) Missing’, and the focus of this paper is imputation of such ‘Temporarily Missing’ prices.

The overall mean imputation methodology as detailed in IMF’s CPI Manual, 2020, (Graf B, 2020), is being adopted in most countries by their NSOs, for imputation of ‘Temporarily Missing’ prices. The aim of this paper is to assess and compare the performance of, the conventional overall mean imputation methodology, that is adopted by most NSOs, referred to here as ‘OMI’, to a more advanced and model based imputation technique ‘Multiple Imputation (MI)’ methodology, for compilation of the component consisting of ‘Food-Items’ of CPI.

2.1 Missing values – what are they and why are they important:

Simply said, missing data are observations that were meant to be measured but were not because of their inevitable nature. Missing observations often result in incomplete data sets, a problem that almost every research study faces. It is often said that one should be surprised if no data is missing. However, the problem is not taken as serious by investigators or researchers because they do not understand the implications of the missing data on the final result. However, a number of researchers have been working in this area and have been proposing different analytical methods to handle this problem. The literature available on the subject has been growing and reflects the fact that the last word, on incomplete data issues, is yet to be pronounced.

2.2 Need for missing data analysis:

While dealing with missing values it is often assumed that a small number of rows with missing entries in the data matrix (Rencher, 2002), does not constitute a serious problem. Each row that has a missing value is simply discarded. However, with this procedure, if the small portion of missing data is widely distributed, it would lead to a substantial loss of data. For example, in a large data set with sample size, n = 550 and variable size, p= 85, if only about 1.5% of the 550 × 85 = 46,750 measurements were missing, but nearly half of the observation vectors (rows of Y) turned out to be incomplete, as it depends on the spread of the missing values across different rows. The distribution of missing values in a data set is an important consideration. Randomly missing variable values scattered throughout a data matrix are assumed to be less serious than a pattern of missing values that depends to some extent on the values of the missing variables.

In general, the key issue for analysis of missing data is whether the reasons for missing are related to the outcome of interest. When missing data is unrelated to the outcome, the impact is relatively minor and does not overly complicate the analysis. On the other hand, when it is related to the outcome, greater care is required because there is a potential for the bias when individuals with missing data differ in important ways from those with complete data (Fitzmaurice et al., 2008). Accordingly, when data is incomplete, the reasons for the missing must be examined carefully.

3. Review of CPI expert group papers

Patrick K and Matlhatsi M, (2019) reviewed the papers and presentations from meetings of the 'Ottawa expert group on price statistics', where CPI compilers would look to for advice, and found that sparse coverage of discussions on imputation. They noted that earlier, two presenters had outlined the basic methods for imputation but did not advance the discussion. They also observed that, Roh and Becker-Vermeulen (2013) had attempted to analyse different imputation methods, focusing on the use of imputations in the Swiss CPI when substituting permanently missing items. Patrick K and Matlhatsi M, (2019) also assessed and presented the performance of four imputation methods to estimate prices for temporarily missing varieties. The methods assessed to impute the missing prices were; overall mean, targeted mean, carry forward and Time product dummy (TPD) regression method, based on a 25-month dataset from the South African CPI, comprising 1751 varieties, with products ranging from food items, clothing and furniture. The missing prices in this dataset were imputed by overall mean imputation to ensure a complete matrix for analysis. Then ten percent of price observations were deleted at random to create a data set resembling a dataset realistically faced by price statisticians. The study confirmed that the overall mean and targeted mean imputation methods recommended by the IMF's CPI manual and conventionally adopted in statistics offices are the most reliable, as these methods mostly showed lowest deviation in the imputed price and the overall index.

Hillman B, et al., (2022), presented a review of methods for missing price and product churn in scanner and web-scraped data for the calculation of price indices using multilateral methods.

4. Classification and Analyses Methods for Missing Data

4.1 Classification for Missing Data:

The most widely used missing data classification system that evolved in literature was first introduced by Rubin (1976), who specified three distinct missing data types:

(1) Missing Completely at Random (MCAR),

(2) Missing at Random (MAR) and

(3) Missing Not at Random (MNAR).

The missing data types introduced by Rubin (1976), also relate to the extent of bias that the missing data may exert on the statistical analyses. In the spectrum of MCAR to MNAR, the impact of MCAR on bias is likely to be negligible, whereas with MNAR the impact is likely to be greatest.

All of the causes for missing data were further classified into four classes (also called mechanisms), based on the relationship between the missing data mechanism and the missing and observed values (Martin, 2001; Little & Rubin, 2002; Schafer & Graham, 2002; Enders, 2013). It is important to understand the four classes as the solutions to the problems caused by missing data are different for the four classes.

i. The MCAR assumption being the most stringent of the missing data mechanisms, that can be checked using Little's multivariate test for MCAR (Little, 1988), but unfortunately most missing data are not MCAR.

ii. At the opposite end of the spectrum is Non-Ignorable (NI). NI implies that the missing data mechanism is related to the missing values. For example, in a survey people do not want to respond to something very personal about themselves or something that is unpopular. In an income-survey, it is less likely that higher income individuals reveal income related information, as compared to lower income individuals, which makes missing data on income non-ignorable. Here the value of income determines whether it is observed or missing. At the end of the survey, for income related information, if the proportion of observed data is more among low and moderate-income individuals in the sample, than the higher-income individuals, the estimate of mean income would be lower than the actual population mean. Therefore, complete case analysis, that is analysis based on only fully observed cases, can give highly biased results for NI missing data.

iii. In between these two extremes are Missing at Random (MAR) and Covariate Dependent (CD). Both of these classes require that the cause of the missing data is unrelated to the missing values, but may be related to the observed values of other variables. MAR implies that the missing values are related to either observed covariates or response variables, whereas CD implies that the missing values are related only to covariates. In the same income survey, an example for CD missing data could be that missing income data are not dependent on income but related to another completely observed variable, say education. It may be possible that people with higher education, have not responded to the income related information as compared to individuals with less education.

In short MCAR implies that 'Missing mechanism' does not depend on the observed or missing values, MAR implies that 'Missing mechanism' depends on the observed values but not on the missing values, and MNAR implies that 'Missing mechanism' is dependent on both the observed and missing values.

Implications of MCAR, MAR – when to ignore the **Distribution of Missing mechanism** (DOM)

I. For frequentist statistical procedures, one may ignore the DOM only when the missing data are MCAR
II. For likelihood or *Bayesian* procedures, one may ignore the DOM when the missing data are MAR

Based on implication II, the terminology

$$\text{MAR} \Leftrightarrow \text{ignorable}$$
$$\text{MNAR} \Leftrightarrow \text{nonignorable}$$

The three major mechanisms, MCAR, MAR, and MNAR are not be assumed as mutually exclusive categories of missing mechanisms. In reality, MCAR, pure MAR, and pure MNAR do not exist, as the pure form of any of these missing types are almost unverifiable. Thus the key distinction is whether the mechanism is ignorable (i.e., MCAR, CD, or MAR) or non-ignorable.

4.2 Methods to handle incomplete data

Various techniques exist as a solution to missing data, ranging from data deletion to methods that replace each missing value with an imputed value employing simple statistical analyses methods to techniques based on artificial intelligence. Imputation is the most common solution to handle missing data, where the missing values are estimated and filled in. An important problem of imputation is to preserve the statistical distribution of the data set. This is a complex problem, especially for high-dimensional data. In this paper the description of the missing data analyses methods is restricted to the two methods, OMI and MI, adopted and compared here.

4.3 Imputation Algorithms

Imputation methods may be broadly classified into two categories: random and deterministic (Ming-xiu & Salvucci, 2001). The deterministic imputation approach imputes one and only one possible value to replace each missing value. Once, the imputation scheme is set up, the imputation result is unique. On the other hand, a random imputation method draws imputation values randomly either from the observed data or from the predicted distribution. Multiple sets of imputations can be created to capture the uncertainty between imputations via any random imputation method. Generally, a random imputation method adds more variability to the statistics computed from an imputed data set than a deterministic imputation method.

4.3.1 Overall Mean Imputation

In this method, for a temporarily missing price of an item, the current month price is imputed/derived by multiplying price of the same item in the previous month with average price relative of current month prices to last month prices from rest of markets of the same item where both current and previous month prices are available. The imputation is generally done within town in case of Urban-Index and within state in case of Rural-Index.

The formula in notations is:

$$\boldsymbol{P_{ik(imputed)}^{t} = P_{ik}^{t-1} \times \frac{\sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{P_{ij}^{t}}{P_{ij}^{t-1}}}{m-1}} \qquad \textbf{----------- (3)}$$

Where; $P_{ik(imputed)}^{t}$ *is the imputed price of the ith item missing from the kth market at month t*

P_{ik}^{t-1} *is the price of the ith item recorded from the kth market at month t-1*

P_{ij}^{t} *and* P_{ij}^{t-1} *are the prices of the ith item recorded from the jth market at month t and t-1 respectively*

(m-1) – the total number of markets from where the price for ith item is recorded at month t.

The OMI imputation technique adopted here is as per the detailed illustration, in the international manual – Consumer Price Index Manual: Concepts and Methods, 2020, (Graf B, 2020), brought out by the consortium of IMF, ILO, OECD, etc.. Further, suggested this method as a standard alternative technique, to be adopted for micro-level imputation for temporarily missing item price data in regular compilation of CPI.

4.3.2 Multiple Imputation (MI)

The inherent bias of deterministic imputation may be eliminated through random imputation. The standard error estimates that result from treating the deterministic imputed data, as though it were real data, are typically too low, and test statistics are excessively high. Whereas, when using random imputation, several ‘completed’ data sets are produced by repeatedly performing the imputation process. The estimates of the parameters of interest would vary slightly for every imputed data set due to the random component. This variability across imputations can be used to adjust the standard errors upward (Allison, 2001).

The MI method involves imputing m values for each missing cell in the data matrix and creating m ‘completed’ data sets. The observed values are consistent across these completed data sets, but the missing data are replaced by a distribution of imputations representing the uncertainty surrounding the missing values. The m sets of imputations account for the uncertainty about the true values of the missing data. After the multiple imputations are created, m plausible versions of the complete-data exist, each of which are analysed by standard complete data methods. The results of the m analyses are then combined to produce a single inferential statement that includes uncertainty due to missing data (Schafer, 1997).

For proper imputation the primary requirement is that the coefficients of the imputation model must be (nearly) unbiased and consistent, and that the specification of the imputation model must be consistent with the posited mechanism of missing. In practice, this means (i) that the imputation model must be a ‘good’ model for predicting missing mechanism, and (ii) in case there is any association between the variable with missing data (X) and the outcome variable (Y) of the substantive model, then Y must be included in the imputation model.

Capturing the variability in the estimated parameters of the imputation model is the other prerequisite for proper imputation. For example, frequent hotdeck draws do not represent ‘proper’ imputation because they only capture sample level uncertainty about the missing data, and not the population level uncertainty. A proper imputation model must be structured to account for the variability in parameter estimates that would come from different samples drawn from the population that is implicit in the imputation of the missing data.

4.4 Multiple Imputation Software

Amelia II – is based on Bootstrap plus Expectation-Maximization algorithm to impute missing values from a dataset and produces multiple output datasets for analysis. Also 'Amelia' has a provision to indicate that the data is a time-series data and to indicate in its imputation syntax the column or variable name identifying time in time series data.

MICE – does multiple imputations based on the algorithm - Multivariate Imputation by Chained Equations. R software includes both packages 'Amelia' and 'MICE', among others for implementing MI for missing data. However, in this paper for the purpose of analysis 'MICE' algorithm was adopted as it more advanced, with literature on its advanced computing applications only growing.

5. Methodology and Plan of analysis

For compilation of the component consisting of 'Food-Items' of the CPI, the steps illustrated under section 2, (a) to (c) are carried out and further procedures as illustrated here below were adopted:

(d) Applying the procedure (a) to (c), the compilation of the component consisting of 'Food-Items' of Consumer Price Index was done for the complete data as well as the imputed data by the two methods – Overall Mean Imputation (OMI) and Multiple Imputation (MI) by MICE, for MCAR and MAR missing mechanisms with 5, 10 and 15 percent missing values.

(e) Since the focus of this paper is imputation of 'Temporarily Missing' prices, examining the reasons for such missing data types, the missing data to be 'atleast MAR if not MCAR' is a valid assumption and therefore one of the requirements for proper imputation being missing mechanism to be 'Ignorable' is met.

5.1 Weighting diagram for the analysis

The original weights of the item categories as per the CPI-manual-NSO India, (CPI, 2015), was proportioned to 100 as per the individual food item categories data available for analysis. The weighting diagram considered here for the index compilation for 'Item Categories' is as per the below table;

Table 1 – Weighting diagram

Food Item Category	Weights as per CPI Manual*	Weights adopted	Number of Items considered
Cereals Products	6.58701	26.61	3 (Rice, Wheat, *Atta*)
Milk Products	5.32597	21.52	1 (Milk)
Non Alcoholic beverages	1.13066	4.57	1 (Tea)
Oils Fats	2.81093	11.36	6 (Oils of Groundnut, Mustard, Palm, Soya, Sunflower and *Vanaspati)*
Pulses Products	1.72847	6.98	5 (*Dals* of *Arhar*, Gram, *Masoor, Moong, Urad*)
Spices	1.79120	7.24	1 (Salt)
Sugar & Confectionary	0.97201	3.93	2 (Sugar and *Gur*)
Vegetable	4.40768	17.81	3 (Onion, Tomato and Potato)
Total	**24.75393**	**100**	**22**

**(CPI, 2015-*https://Cpi.Mospi.Gov.In/Pdfile/Cpi-Changes_In_The_Revised_Series.Pdf, Page 69*)*

5.2 Empirical Data

In India, for CPI compilation, the NSO collects the monthly price data, from 1181 villages and 1114 markets in 310 selected towns (CPI, 2015). This unit level price data is not available on public domain. For the purpose of this paper, to be able to use the NSO's weighting diagram of the current series adopted for the compilation of CPI in India, the data was required in the same structure. As it is difficult to obtain the similar time-series data across all item categories, the index compilation and comparison across the two imputation methods with the complete data index, is restricted to the food items component of the CPI.

Hence, the author explored the public domain for sources of unit level price data of the required food item basket. The data suitable for the required analysis was found on the website of the 'Price Monitoring Division', of the Department of Consumer Affairs in the Ministry of Consumer Affairs, of Government of India.

Initially considered, the 60 months data, reported every second Monday for the months falling in the period January 2018 to December 2022, (for non-seasonal 22 food items that fall under the eight food item categories). For the analysis, considered data of only 51 urban centres that had complete data on all 22 food items, as the objective of this paper is to compare and demonstrate the alternative missing data imputation techniques,. However, it was found that this data had limited variation over time and the price index of food items were concentrated around the mean value of 101.2, for almost all 60 months. Hence, adopted statistical simulation to generate the required empirical data.

5.3 Data Simulation

Precisely, considered the prices of the second Monday of January 2018 as reported on the website for the base-line price data (22 food items that fall under 8 food item categories, across 51 urban centres, from 18 states). Then considered subsequent 23 months data from the same source with the same structure, item-wise, reported for second Monday of every month, (February 2018 to December 2019) for the Mean and Standard Deviation estimates in the statistical simulation plan. However, for standard deviation, an additional variation of 'random number ~ 1 to 1.2' was added. Thus, two column vectors generated; mean and standard deviation of length 1122, based on the 23 months price data reported on the website.

Using the functions available in package of R software, the preliminary price data for the 59 months were simulated and for any negative values generated, replaced with the value - (January 2018 price + random number ~ 1 to 3), using the random number generation functions available in package of R software. Further tested the simulated data for validity of stationarity assumption of time-series data using Augmented Dickey Fuller (ADF) Test and the simulations were found to be a stationary time series with one-month-lag, with 95% confidence. The ADF Test results of one instance of simulated data are included at Table 7 in the Annexure of this paper. Therefore, the complete data considered here, after simulation, was a data matrix consisting of 1122 rows and 59 columns plus the base-line data.

5.4 Generation of missing data

Missing data of 5, 10 and 15 percent were generated for MCAR and MAR missing mechanisms. MCAR and MAR missing data, of the required percent was generated across the 59 months data by specifying the 'mechanisms' as '*mcar*' and '*mar*' with the functions available in packages of R software.

5.5 Missing Data Imputation

(a) CPI Overall Mean Imputation (OMI)

In the current month, if an item price is 'Temporarily missing', it is imputed by multiplying price of the same item in the previous month with average price relative of current month prices to last month prices for the same item, from the rest of markets, given that, both current and previous month prices of the item are available. The formula in notations is as illustrated above at (3).

Here, using the formula (3), the missing prices for an item in the current month were imputed by multiplying price of the same item in the previous month with average price relative of current month prices to last month prices for the same item from the rest of 'Centres' where both current and previous month prices were available.

(b) Multiple Imputation (MI)

Multiple imputation (MI) of missing price data was done using the *'MICE'* function available in packages of R software, as it has a provision to indicate the data as time-series data and include the time variable in the imputation syntax. The MI by MICE was done with a lag of 1 month, specifying 5 imputations and 5 iterations. The 5 imputed datasets, indicated here as, MICE1 to MICE5 were considered for the MI imputed data set for further analysis and comparison of results, without pooling to be able to assess the performance of individual MI results vis-à-vis OMI.

5.6 Simulation Plan

Simulation was carried out specifying *sample Size* = (50, 100, 250, 500), and *100 replications* for 5 and 10 percent missing, whereas *120 replications* for 15 percent missing, as some of the results indicated 'error of insufficient rows', in 15 percent missing. The detail analysis and results compared in the simulation is included below.

First, the summary statistics for the indices being compared were generated. Then the results across the two methods, i.e., the indices compiled for complete data, OMI and MI data for MCAR and MAR missing mechanisms with 5, 10 and 15 percent missing values, was compared using the Repeated Measures Analysis of Variance (RM-ANOVA) test procedure. The RM-ANOVA requires that the data being compared satisfy the assumptions of – (1) data is continuous, (2) data has no outliers (3) data is from normal distribution. Here the data being food items price indices, are continues. The data was checked for outliers and as data being compared consisted of 780 observations for each of 5, 10 and 15 percent missing data, (60 months index data for each of - Complete, OMI_MCAR, OMI_MAR, MICE1_MCAR, MICE2_MCAR, MICE3_MCAR, MICE4_MCAR, MICE5_MCAR, MICE1_MAR, MICE2_MAR, MICE3_MAR, MICE4_MAR, MICE5_MAR) the normal distribution of the data was assumed by 'Law of Large Numbers'. Then the RM-ANOVA test was performed on the indices compiled for complete data, OMI and MI data for MCAR and MAR missing mechanisms with 5, 10 and 15 percent missing values, using the *'anova_test'* function available in package of R software. Further for the significant results the post-hoc *'Bonferroni'* pair-wise comparison tests were also done using the *'pairwise_t_test'* available in package of R software.

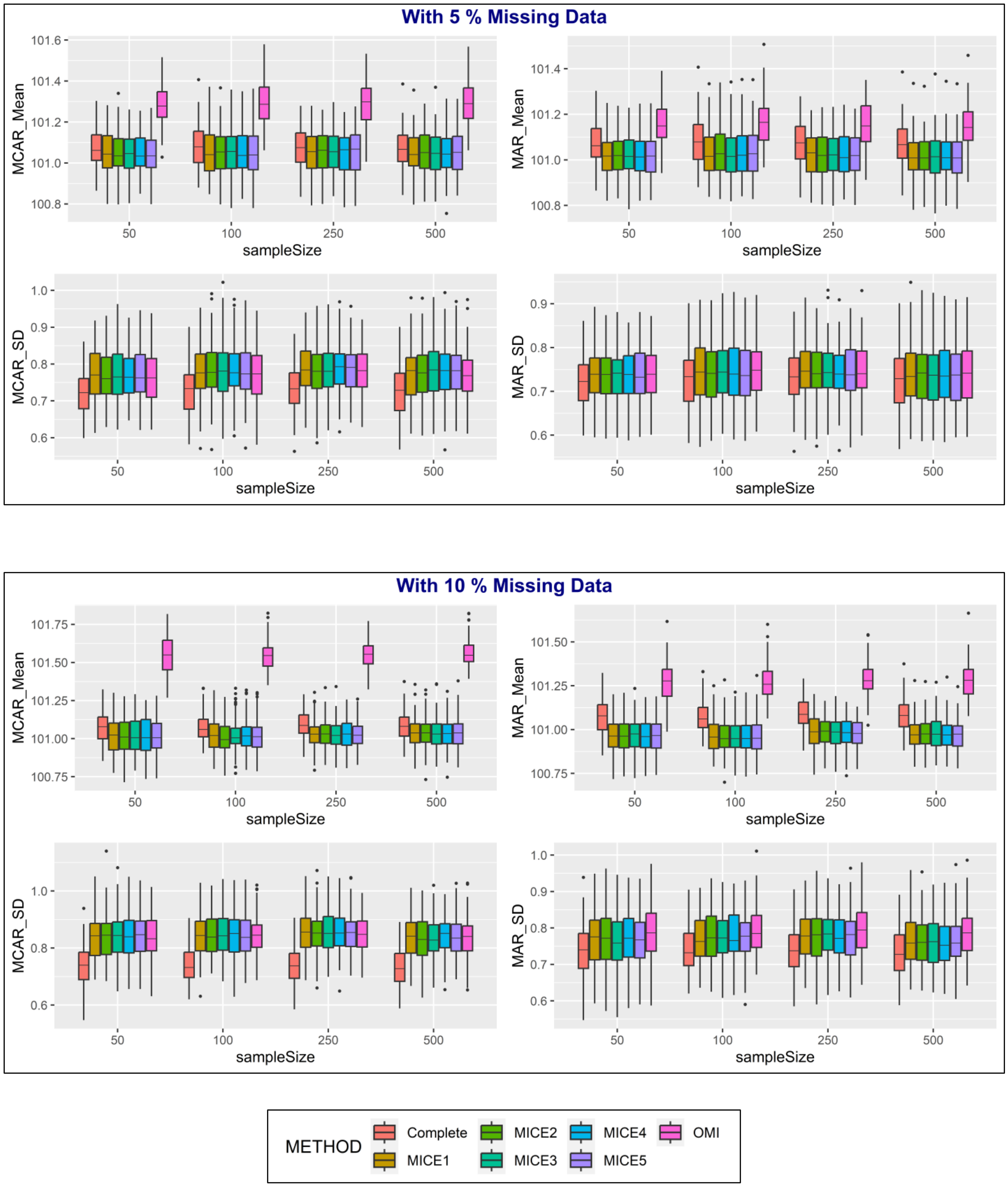
The simulation results consisted of 400 observations each for 5 and 10 percent missing and 422 observations for 15 percent missing, across MCAR and MAR data.

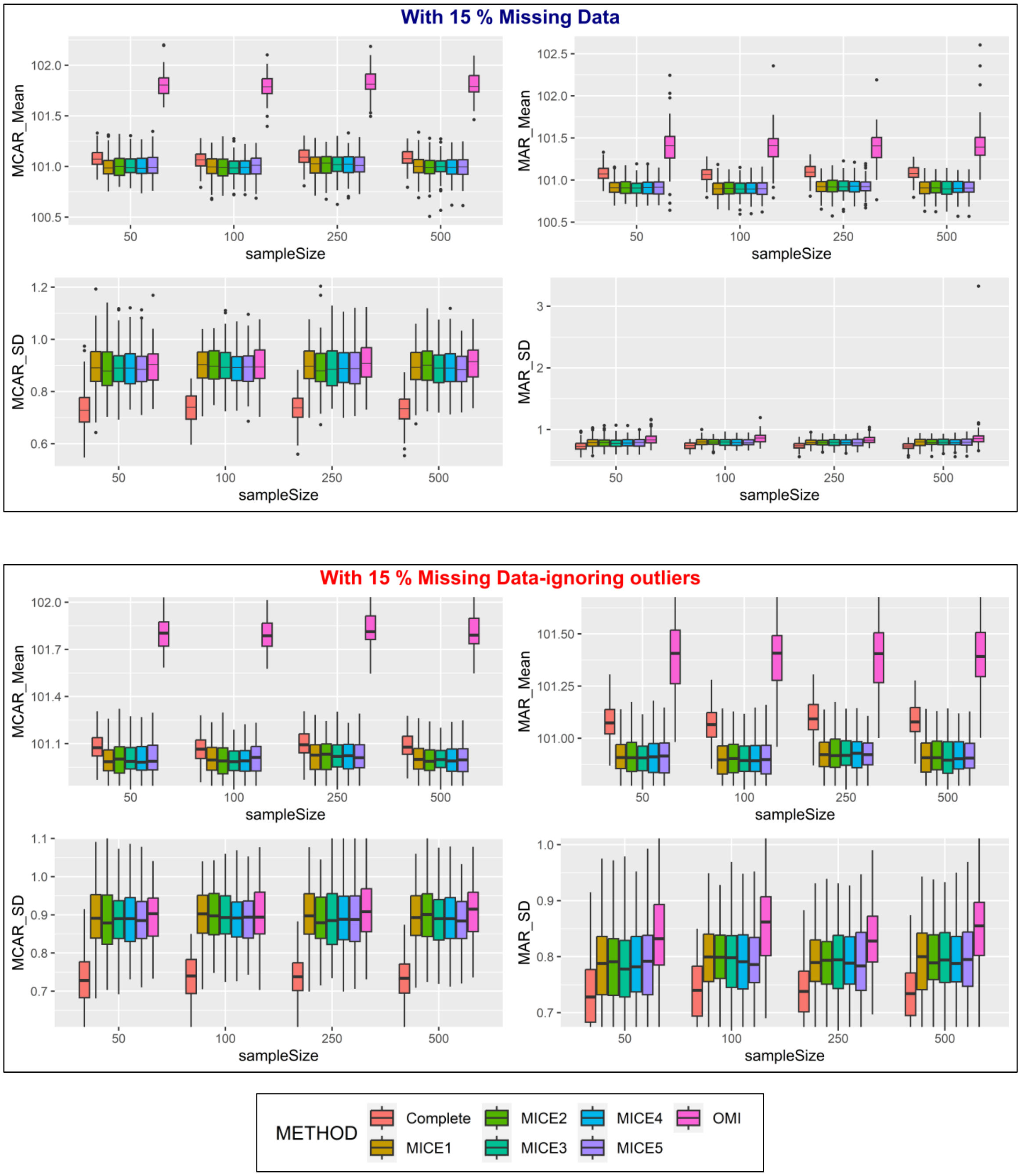
6. Results

6.1 Summary Statistics

Detailed summary results are included at Table 4, 5 and 6 in the Annexure of this paper, while graphical representation of summary results is included here. From the 'Figure 1', below it is evident that for 5, 10 and 15 percent MCAR, the mean vector estimates based on MI, for all 5 imputations, MICE1 through MICE5, is close to the complete data mean vector across all sample sizes, 50, 100, 250 and 500. Further, for MAR this deviation is minimal in estimated mean vectors for 5, 10 and 15 percent from the complete data mean vector, for all 5 imputations, MICE1 through MICE5. While the mean vector estimates based on the OMI method, is higher than the actual mean computed by complete data mean vector across all sample sizes, 50, 100, 250 and 500, and across all percentages, 5, 10 and 15, missing, for MCAR as well as MAR. The SD vector by both methods across all sample sizes, missing percentages and mechanisms hovers around the 0.7~0.9 slightly deviating from the complete data SD vector. Particularly for 15% MCAR and MAR, with OMI imputations the presence of outliers in mean as well as standard deviations is clearly visible. Therefore, the additional graphs 'ignoring-outliers' is included for 15% missing data.

Figure 1 – Comparison of Mean and Standard Deviations of the complete data with estimates of imputed data





6.2 Hypothesis tested by Repeated Measures – ANOVA (RM-ANOVA)

Null Hypothesis, H_0 – The means of the computed index for the imputed data sets by Overall Mean Imputation (OMI) as well as Multiple Imputation (MI) Method is the same as mean of the computed index for the complete data across 5%, 10% and 15% MCAR and MAR.

$$H_0: \mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$$

Against the alternate hypothesis,

H_A: at least two means are significantly different

Table 2 – Simulations significant across all sample sizes, 50, 100, 250 and 500

Missing data %	Total Rows	Post-hoc test recommended for rows significant at 5 %
5	400	400
10	400	400
15	422	422

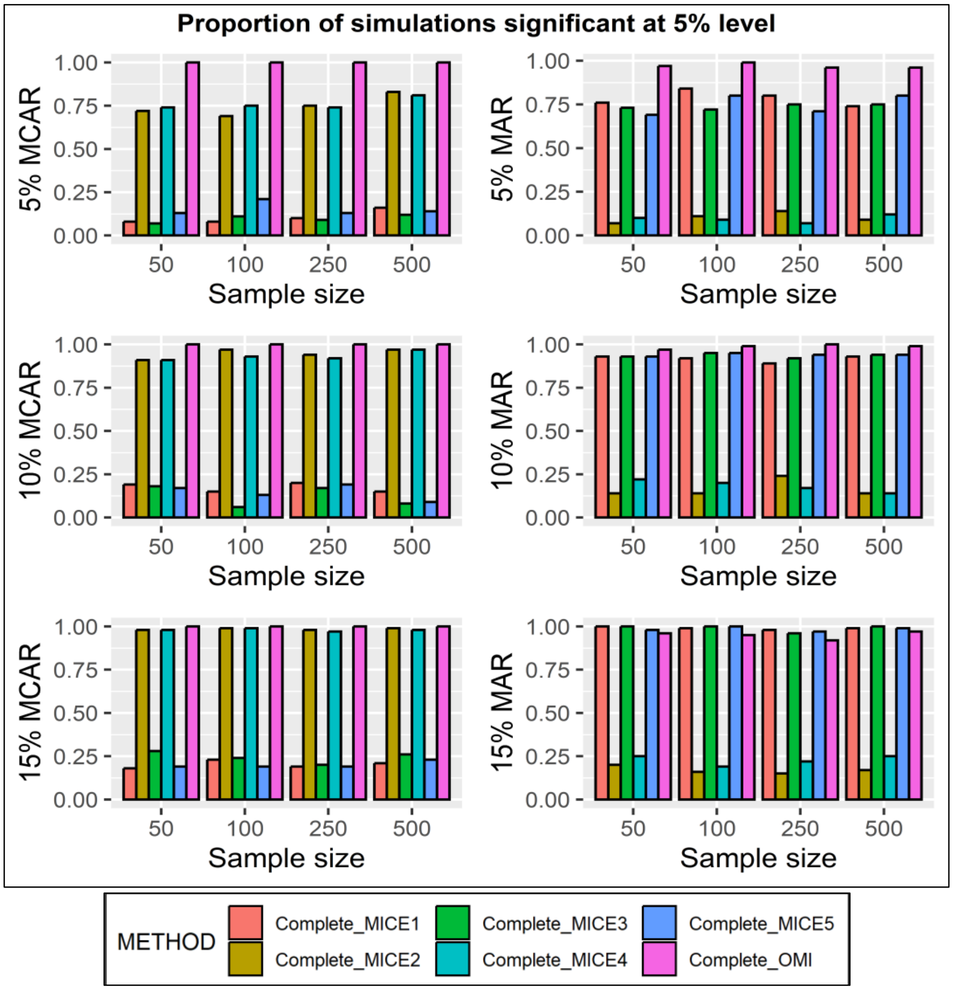
Therefore, it was concluded that H_0 may be rejected for all rows of simulation and '*bonferroni*' post-hoc pair-wise comparison test was carried out to further explore the performance of each of the imputation indices with the complete data indices.

6.3 Results of Post-hoc 'B*onferroni*' pair-wise comparison tests

Table 3 – Proportion of simulations significant at 5 percent level, in pair-wise comparison of mean vectors across sample sizes

N *Sample Size*	**COMPLETE vis-à-vis**	**5 % Missing**		**10 % Missing**		**15 % Missing**	
		MCAR	**MAR**	**MCAR**	**MAR**	**MCAR**	**MAR**
50	**MICE1**	0.08	0.76	0.19	0.93	0.18	1.00
	MICE2	0.72	0.07	0.91	0.14	0.98	0.20
	MICE3	0.07	0.73	0.18	0.93	0.28	1.00
	MICE4	0.74	0.10	0.91	0.22	0.98	0.25
	MICE5	0.13	0.69	0.17	0.93	0.19	0.98
	OMI	1.00	0.97	1.00	0.97	1.00	0.96
100	**MICE1**	0.08	0.84	0.15	0.92	0.23	0.99
	MICE2	0.69	0.11	0.97	0.14	0.99	0.16
	MICE3	0.11	0.72	0.06	0.95	0.24	1.00
	MICE4	0.75	0.09	0.93	0.20	0.99	0.19
	MICE5	0.21	0.80	0.13	0.95	0.19	1.00
	OMI	1.00	0.99	1.00	0.99	1.00	0.95
250	**MICE1**	0.10	0.80	0.20	0.89	0.19	0.98
	MICE2	0.75	0.14	0.94	0.24	0.98	0.15
	MICE3	0.09	0.75	0.17	0.92	0.20	0.96
	MICE4	0.74	0.07	0.92	0.17	0.97	0.22
	MICE5	0.13	0.71	0.19	0.94	0.19	0.97
	OMI	1.00	0.96	1.00	1.00	1.00	0.92
500	**MICE1**	0.16	0.74	0.15	0.93	0.21	0.99
	MICE2	0.83	0.09	0.97	0.14	0.99	0.17
	MICE3	0.12	0.75	0.08	0.94	0.26	1.00
	MICE4	0.81	0.12	0.97	0.14	0.98	0.25
	MICE5	0.14	0.80	0.09	0.94	0.23	0.99
	OMI	1.00	0.96	1.00	0.99	1.00	0.97

Figure 2 – Proportion of simulations significant across MICE1 to MICE5 and OMI in Post-hoc '*Bonferroni*' pair-wise-comparison of means with complete data



The Results of Post-hoc '*Bonferroni*' pair-wise comparison tests from the Figure 2 and Table 3 clearly indicate – the number of simulation results significant is only about 10 to 20 percent with MCAR data, across 5, 10 and 15 percent missing for atleast 3 of the 5 MIs. On similar lines, with MAR data, across 5, 10 and 15 percent missing, the number of simulation results significant is only about 10 to 20 percent, for atleast 2 of the 5 MIs. While, with OMI it is evident that the number of simulation results significant is 92 percent to 100 percent, for MCAR and MAR, across 5, 10 and 15 percent missing. The simulation results clearly show the better performance of MIs over OMI. The performance of MI could be further improved by increasing the number of iterations and imputations. Here for demonstration purpose the iterations as well as imputations were set at five. Given that, van Buuren, S (2018), has found, imputing a dataset in practice often involves trial and error to adapt and refine the imputation model and that in such initial explorations one does not require large m. It would be convenient to set m=5 during model building, and increase m only when a satisfied model for the final round of imputation is obtained. Further, pooling of multiple imputations was not carried, to be able to compare the performance of each of the five imputations in each scenario directly with the OMI.

7. Summary and Conclusion

A comparison of CPI for food items component, compiled for complete data, imputed data, by Overall Mean Imputation (OMI) and Multiple Imputation (MI), was done based on RM-ANOVA test followed by post-hoc '*Bonferroni*' pair-wise comparison tests. Simulations were carried for sample sizes = c (50, 100, 250, 500), for both MCAR and MAR, with 100 replications, for 5 and 10 percent missing, 120 replications, for 15 percent missing. The test results clearly show the price index compilation with Multiple Imputation for the missing price data were closer to the complete data price indices. With less than 10 percent, significant, for atleast 3 MIs, for MCAR, and for MAR, less than 20 percent significant, for atleast 2 MIs. While it was more than 92 percent significant for OMI imputed data, while comparing the means with complete data, in pair-wise Bonferroni test.

While working with the actual unit level data, the price statistician may expect the missing data to be much less than 5 percent and believe that there is no harm in carrying on with the OMI, as it is the international conventional method. However, at the end of the day the OMI method is a deterministic imputation technique and does not take into account the distribution of observed data while imputing the missing data. Therefore, as an alternative imputation method, that is random as well as takes care of the distribution of observed data while imputing the missing data, 'Multiple Imputation', needs to be explored and adopted in future by the NSOs. In reality the variety of items, specifications as well as locations of each food item are much more in number and therefore, with the actual unit level data the results of MI in comparison of OMI, may further improve due to greater volume of data across all items.

Future research needs to be carried out by increasing the number of iterations as well as imputation to greater than five, through parallel computing techniques and with an advanced methodology, for pooling the imputed index number data from the multiple imputations.

8. Limitation:

The unit level data required for compilation of CPI is a time-series price data of item basket included for the particular series. In addition, this data should be available in the specific structure, as per the weighting diagram.

NSO's CPI time-series unit level data is not available on the public domain *(as per government policy)*. Therefore, the author explored alternative data sources for empirical data to be considered for the analysis in this paper. The online data available on the website of Department of Consumer Affairs in the Ministry of Consumer Affairs, of Government of India, was found to be suitable for analysis.

Precisely, considered the prices reported for second Monday of January 2018 as reported on the website for the base-line price data (22 food items that fall under 8 food item categories, across 51 urban centres, from 18 states). Then considered subsequent 23 months data from the same source with the same structure, item-wise, reported for second Monday of every month, (February 2018 to December 2019) for the Mean and Standard Deviation estimates in the statistical simulation plan.

Acknowledgements:

The author would like to acknowledge the Coordination and Quality Control Division and the anonymous Referees for providing their valuable feedback and comments, which helped in improving the quality of the paper. The author would also like to thank Prof. D K Subbakrishna, former Professor and Head, Department of Biostatistics, NIMHANS, Bangalore for his advice on improving the manuscript.

References:

1. Allison Paul D (2001): Missing Data. Iowa City: Sage Publications.
2. CPI (2015), Consumer Price Index, Changes in the Revised Series, 2015 - MOSPI, CSO, NAD, Prices and Cost of Living Unit, Available at https://Cpi.Mospi.Gov.In/Pdfile/Cpi-Changes_In_The_Revised_Series.Pdf (accessed 25 February 2023).
3. Enders, C. K. (2013). Analyzing structural equation models with missing data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course*, IAP, (pp. 493–519).
4. Graf B (2020) *Consumer Price Index Manual, 2020: Concepts and Methods*: Washington, DC: International Monetary Fund. Available at: https://www.elibrary.imf.org/display/book/9781484354841/9781484354841.xml (accessed 1 March 2023).
5. Hillman B, Greenhough L, Rose A, Sands H, (2022), paper presentation; What do missing prices mean for the choice of index number method with alternative data?, Seventeenth Meeting of the Ottawa Group, International Working Group On Price Indices, - Rome, 07-10 June 2022, available at: https://www.ottawagroup.org/ottawa/ottawagroup.nsf/4a256353001af3ed4b2562bb00121564/c7e600e0bec85d82ca25888400059a93?OpenDocument#Meeting%2017%20Rome, (Accessed on 6 December, 2023)
6. Little R J A & Rubin DB (1987, 2002) Statistical Analysis with Missing Data. First & Second Edition John Wiley & Sons: New York
7. Little R J A (1988) : A test of missing completely at random for multivariate data with missing values; Journal of the American Statistical Association, 83, 1198-1202
8. Martin K G (2001): Missing Data Mechanisms; StatNews #46, October 9
9. Patrick K & Matlhatsi M, (2019): Statistics South Africa, paper presentation; Missing in action: testing alternative imputation methods in price statistic, Sixteenth Meeting of the Ottawa Group, International Working Group On Price Indices, - Rio de Janeiro, Brazil 08-10 May 2019, available at: https://www.ottawagroup.org/Ottawa/ottawagroup.nsf/home/Meeting+16/$FILE/Missing%20in%20action%20 paper.pdf (accessed on 6 December, 2023)
10. Rencher Alvin C, (2002): Methods of multivariate analysis, second edition, Wiley series in probability and mathematical statistics, 2002 by John Wiley & Sons

11. Report of the Group of Technical Advisory Committee on Statistics of Prices and Cost of Living, (2014) - MOSPI, CSO, NAD, Prices and Cost of Living Unit, Available at: https://cpi.mospi.gov.in/PDFile/Report_of_the_Group_of_TAC_on_SPCL_on_Base_Revision_of_CPI.pdf (accessed 25 February 2023).
12. Roh, C & Bekker-Vermeulen, C, (2013): Statistics Norway, paper presentation; Clothing: the use of class mean imputation in the Swiss Consumer price index (CPI) – analysis and impact on the results, Thirteenth Meeting of the Ottawa Group, International Working Group On Price Indices,- Copenhagen, Denmark, 01-03 May 2013, available at: https://www.ottawagroup.org/Ottawa/ottawagroup.nsf/4a256353001af3ed4b2562b00121564/8bdac0e73d96c891ca257bb00002fdb4/$FILE/Corinne%20Becker%20Bridged%20Overlap%20in%20the%20Swiss%20CPI_Ottawagroup2013.pdf (accessed on 6 December, 2023)
13. Rubin, D. B. (1976): Inference and Missing Data; Biometrika, 63, 581– 590
14. Schafer, J. L. (1997): Analysis of Incomplete Multivariate Data, New York: Chapman & Hall.
15. Schafer J L. and Graham John W. (2002): Missing Data: Our View of the State of the Art - Psychological Methods, 2002, Vol. 7, No. 2, 147–177
16. van Buuren, S (2018); Flexible Imputation of Missing Data. Second Edition, CRC Press, Boca Raton, FL, Available at: https://stefvanbuuren.name/publication/vanbuuren-2018/ (accessed 8 April 2023).
17. Website of Price Monitoring Division of the Department of Consumer Affairs in Ministry of Consumer Affairs, Food and Public Distribution, Available at: https://consumeraffairs.nic.in/price-monitoring-cell/price-monitoring-cell (accessed 26 February 2023).

Annexure

Summary results

Table 4: Food Indices computed for the Complete and imputed data with 5% Missing

N (Sample Size)	METHOD	MCAR						MAR					
		MIN		MAX		MEAN		MIN		MAX		MEAN	
		MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
50	Complete	100.87	0.599	101.30	0.861	101.07	0.722	100.87	0.599	101.30	0.861	101.07	0.722
	MICE1	100.80	0.613	101.28	0.918	101.05	0.772	100.82	0.595	101.25	0.893	101.02	0.736
	MICE2	100.80	0.629	101.34	0.931	101.05	0.770	100.83	0.592	101.24	0.874	101.02	0.736
	MICE3	100.80	0.622	101.26	0.963	101.05	0.769	100.78	0.594	101.23	0.881	101.02	0.737
	MICE4	100.85	0.647	101.26	0.926	101.05	0.770	100.82	0.588	101.25	0.857	101.02	0.736
	MICE5	100.80	0.621	101.27	0.946	101.04	0.771	100.82	0.596	101.25	0.881	101.02	0.737
	OMI	101.03	0.622	101.52	0.938	101.28	0.765	100.94	0.601	101.39	0.872	101.15	0.740
100	Complete	100.88	0.582	101.41	0.901	101.08	0.734	100.88	0.582	101.41	0.901	101.08	0.734
	MICE1	100.85	0.571	101.37	0.953	101.06	0.780	100.84	0.573	101.33	0.909	101.03	0.748
	MICE2	100.80	0.568	101.37	0.991	101.06	0.784	100.83	0.587	101.34	0.908	101.03	0.744
	MICE3	100.78	0.597	101.36	1.022	101.06	0.785	100.82	0.603	101.34	0.924	101.03	0.746
	MICE4	100.83	0.605	101.35	0.976	101.06	0.784	100.84	0.590	101.35	0.927	101.03	0.746
	MICE5	100.78	0.572	101.36	0.973	101.05	0.785	100.83	0.587	101.35	0.914	101.03	0.746
	OMI	101.06	0.581	101.58	0.945	101.29	0.778	100.97	0.608	101.51	0.920	101.16	0.750
250	Complete	100.84	0.563	101.28	0.882	101.07	0.733	100.84	0.563	101.28	0.882	101.07	0.733
	MICE1	100.79	0.627	101.28	0.940	101.05	0.783	100.81	0.589	101.22	0.914	101.02	0.747
	MICE2	100.80	0.586	101.26	0.958	101.05	0.781	100.80	0.575	101.23	0.890	101.02	0.746
	MICE3	100.84	0.620	101.30	0.962	101.05	0.781	100.80	0.601	101.23	0.931	101.02	0.747
	MICE4	100.78	0.616	101.25	0.969	101.05	0.788	100.83	0.565	101.24	0.909	101.02	0.744
	MICE5	100.79	0.641	101.28	0.957	101.05	0.786	100.80	0.572	101.23	0.890	101.02	0.745
	OMI	101.01	0.629	101.53	0.921	101.29	0.780	100.91	0.599	101.35	0.930	101.15	0.748
500	Complete	100.84	0.568	101.39	0.901	101.07	0.725	100.84	0.568	101.39	0.901	101.07	0.725
	MICE1	100.80	0.611	101.36	0.980	101.05	0.777	100.78	0.591	101.34	0.949	101.01	0.738
	MICE2	100.81	0.606	101.29	0.979	101.05	0.778	100.79	0.586	101.32	0.931	101.01	0.740
	MICE3	100.81	0.610	101.37	0.982	101.04	0.778	100.77	0.588	101.38	0.925	101.01	0.736
	MICE4	100.75	0.567	101.31	0.994	101.04	0.775	100.80	0.584	101.35	0.918	101.01	0.739
	MICE5	100.84	0.618	101.31	0.970	101.05	0.777	100.79	0.595	101.33	0.910	101.01	0.735
	OMI	101.06	0.611	101.57	0.975	101.29	0.770	100.90	0.596	101.46	0.915	101.15	0.740

Table 5: Food Indices computed for the Complete and imputed data with 10% Missing													
N (sample_size)	**METHOD**	**MCAR**						**MAR**					
		MIN		**MAX**		**MEAN**		**MIN**		**MAX**		**MEAN**	
		MEAN	**SD**	**MEAN**	**SD**	**MEAN**	**SD**	**MEAN**	**SD**	**MEAN**	**SD**	**MEAN**	**SD**
50	**Complete**	100.85	0.547	101.32	0.939	101.07	0.739	100.85	0.547	101.32	0.939	101.07	0.739
	MICE1	100.78	0.689	101.30	1.051	101.02	0.839	100.72	0.593	101.20	0.949	100.96	0.770
	MICE2	100.71	0.684	101.28	1.140	101.01	0.839	100.73	0.572	101.21	0.963	100.96	0.771
	MICE3	100.79	0.648	101.29	1.082	101.02	0.842	100.72	0.555	101.24	0.946	100.97	0.762
	MICE4	100.74	0.656	101.25	1.050	101.02	0.842	100.74	0.580	101.19	0.938	100.96	0.771
	MICE5	100.74	0.656	101.28	1.037	101.01	0.841	100.74	0.590	101.19	0.935	100.96	0.768
	OMI	101.27	0.631	101.82	1.014	101.54	0.839	100.99	0.587	101.62	0.976	101.27	0.790
100	**Complete**	100.90	0.620	101.33	0.905	101.07	0.744	100.90	0.620	101.33	0.905	101.07	0.744
	MICE1	100.80	0.631	101.32	1.029	101.03	0.847	100.79	0.636	101.25	0.910	100.97	0.772
	MICE2	100.76	0.711	101.27	1.019	101.01	0.846	100.70	0.625	101.28	0.936	100.96	0.775
	MICE3	100.77	0.683	101.33	1.042	101.02	0.845	100.74	0.608	101.21	0.926	100.96	0.778
	MICE4	100.80	0.629	101.32	1.038	101.03	0.847	100.73	0.616	101.21	0.922	100.96	0.777
	MICE5	100.79	0.677	101.30	1.040	101.02	0.846	100.74	0.590	101.31	0.930	100.96	0.779
	OMI	101.35	0.687	101.82	1.021	101.54	0.846	101.06	0.672	101.60	1.011	101.27	0.794
250	**Complete**	100.88	0.585	101.29	0.906	101.09	0.740	100.88	0.585	101.29	0.906	101.09	0.740
	MICE1	100.79	0.687	101.30	1.052	101.03	0.859	100.74	0.635	101.20	0.930	100.98	0.776
	MICE2	100.83	0.660	101.34	1.072	101.03	0.855	100.78	0.590	101.16	0.957	100.98	0.777
	MICE3	100.81	0.699	101.34	1.051	101.03	0.854	100.76	0.616	101.19	0.936	100.98	0.781
	MICE4	100.81	0.649	101.26	1.045	101.04	0.858	100.74	0.626	101.16	0.920	100.98	0.777
	MICE5	100.83	0.703	101.26	1.047	101.03	0.855	100.78	0.609	101.13	0.964	100.98	0.778
	OMI	101.32	0.696	101.77	0.993	101.56	0.850	101.03	0.644	101.54	0.980	101.29	0.796
500	**Complete**	100.88	0.588	101.38	0.891	101.09	0.731	100.88	0.588	101.38	0.891	101.09	0.731
	MICE1	100.80	0.667	101.36	1.011	101.04	0.835	100.79	0.631	101.28	0.959	100.98	0.763
	MICE2	100.73	0.626	101.32	1.002	101.04	0.832	100.78	0.628	101.27	0.954	100.98	0.766
	MICE3	100.82	0.661	101.36	1.020	101.03	0.831	100.80	0.623	101.27	0.919	100.97	0.763
	MICE4	100.75	0.654	101.31	0.989	101.04	0.844	100.79	0.619	101.30	0.924	100.97	0.759
	MICE5	100.81	0.690	101.38	1.027	101.04	0.838	100.78	0.605	101.25	0.974	100.97	0.765
	OMI	101.39	0.653	101.82	1.028	101.57	0.837	101.08	0.642	101.66	0.986	101.28	0.784

Table 6: Food Indices computed for the Complete and imputed data with 15% Missing													
N (sample_size)	**METHOD**	**MCAR**						**MAR**					
		MIN		**MAX**		**MEAN**		**MIN**		**MAX**		**MEAN**	
		MEAN	**SD**	**MEAN**	**SD**	**MEAN**	**SD**	**MEAN**	**SD**	**MEAN**	**SD**	**MEAN**	**SD**
50	**Complete**	100.87	0.547	101.33	0.974	101.08	0.730	100.87	0.547	101.33	0.974	101.08	0.730
	MICE1	100.75	0.643	101.31	1.193	101.00	0.896	100.70	0.575	101.15	1.031	100.91	0.785
	MICE2	100.80	0.703	101.32	1.141	101.00	0.892	100.72	0.598	101.17	1.067	100.91	0.790
	MICE3	100.79	0.692	101.31	1.118	101.01	0.894	100.68	0.597	101.19	1.071	100.91	0.786
	MICE4	100.74	0.731	101.27	1.121	101.00	0.893	100.70	0.590	101.19	1.067	100.92	0.787
	MICE5	100.76	0.710	101.35	1.113	101.01	0.886	100.70	0.593	101.15	1.000	100.91	0.786
	OMI	101.58	0.733	102.20	1.169	101.80	0.897	100.64	0.664	102.24	1.165	101.39	0.850
100	**Complete**	100.79	0.596	101.28	0.850	101.07	0.739	100.79	0.596	101.28	0.850	101.07	0.739
	MICE1	100.68	0.705	101.24	1.040	101.00	0.897	100.65	0.672	101.19	1.003	100.90	0.797
	MICE2	100.71	0.748	101.30	1.043	100.99	0.900	100.64	0.625	101.13	0.928	100.90	0.797
	MICE3	100.72	0.724	101.28	1.111	100.99	0.902	100.59	0.665	101.15	0.969	100.90	0.795
	MICE4	100.72	0.726	101.22	1.069	100.99	0.889	100.60	0.658	101.15	0.948	100.90	0.794
	MICE5	100.69	0.686	101.23	1.096	101.00	0.892	100.62	0.660	101.16	0.952	100.90	0.795
	OMI	101.40	0.703	102.10	1.077	101.79	0.899	100.79	0.690	102.36	1.194	101.39	0.861
250	**Complete**	100.81	0.560	101.31	0.883	101.10	0.736	100.81	0.560	101.31	0.883	101.10	0.736
	MICE1	100.71	0.699	101.28	1.077	101.02	0.898	100.66	0.648	101.14	0.958	100.92	0.791
	MICE2	100.68	0.673	101.24	1.204	101.02	0.892	100.57	0.634	101.17	0.939	100.93	0.787
	MICE3	100.63	0.734	101.30	1.130	101.02	0.893	100.63	0.625	101.23	0.931	100.93	0.790
	MICE4	100.69	0.699	101.33	1.106	101.02	0.891	100.69	0.616	101.21	0.927	100.93	0.792
	MICE5	100.73	0.706	101.29	1.121	101.02	0.893	100.67	0.630	101.20	0.947	100.92	0.786
	OMI	101.50	0.731	102.19	1.124	101.82	0.912	100.77	0.697	102.19	1.038	101.39	0.836
500	**Complete**	100.80	0.554	101.28	0.874	101.08	0.736	100.80	0.554	101.28	0.874	101.08	0.736
	MICE1	100.69	0.676	101.34	1.060	101.01	0.897	100.63	0.601	101.14	0.943	100.91	0.793
	MICE2	100.51	0.724	101.28	1.119	101.00	0.898	100.62	0.564	101.13	0.938	100.91	0.792
	MICE3	100.57	0.719	101.27	1.076	101.00	0.887	100.62	0.609	101.14	0.933	100.91	0.792
	MICE4	100.62	0.712	101.24	1.119	101.00	0.892	100.57	0.561	101.12	0.950	100.91	0.791
	MICE5	100.61	0.720	101.25	1.033	101.00	0.887	100.57	0.568	101.13	0.969	100.91	0.796
	OMI	101.46	0.736	102.09	1.078	101.81	0.911	101.00	0.657	102.60	3.325	101.42	0.875

Table 7: - Augmented Dickey Fuller Test Results – p-values

(Hypothesis: H_0: The time series is non-stationary – Against the alternate hypothesis – H_A: The time series is stationary with one-lag)

SNO	CENTRE	WHEAT	VANASPATI	URAD_DAL	TOMATO	TEA	SUGAR	SOYA_OIL	SUNFLOWER _OIL	SALT	POTATO	RICE	PALM_OIL	ONION	MUSTARD _OIL	MOONG_ DAL	MILK	MASOOR _DAL	GUR	GROUNDNUT _OIL	GRAM _DAL	ARHAR_ DAL	ATTA
1	ADILABAD	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2	AGRA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.06**	0.01
3	AHMEDABAD	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
4	ALLAHABAD	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01
5	AMBIKAPUR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
6	BHOPAL	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
7	BHUBANESHWAR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
8	BHUJ	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
9	CHENNAI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.05**	0.01	0.01	0.01	0.01	0.01
10	COIMBATORE	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
11	CUTTACK	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
12	DEHRADUN	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01
13	DELHI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.04	0.01	0.01	0.01
14	DINDIGUL	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
15	DURG	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01
16	GORAKHPUR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
17	GURGAON	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
18	GUWAHATI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
19	HALDWANI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
20	HARIDWAR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.01	0.01	0.01	0.01
21	HISAR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
22	INDORE	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
23	JADCHERLA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
24	KANPUR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
25	KARIMNAGAR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01

SNO	CENTRE	WHEAT	VANASPATI	URAD_DAL	TOMATO	TEA	SUGAR	SOYA_OIL	SUNFLOWER _OIL	SALT	POTATO	RICE	PALM_OIL	ONION	MUSTARD _OIL	MOONG_ DAL	MILK	MASOOR _DAL	GUR	GROUNDNUT _OIL	GRAM _DAL	ARHAR_ DAL	ATTA
26	KARNAL	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01
27	KURNOOL	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
28	LUDHIANA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
29	MEERUT	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
30	MUMBAI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
31	NAGPUR	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
32	NASHIK	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
33	PANCHKULA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
34	PATNA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
35	PUNE	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
36	PURNIA	0.01	0.01	0.01	0.01	0.04	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
37	RAIPUR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
38	RAJKOT	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
39	RAMPURHAT	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
40	RANCHI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
41	RUDRAPUR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
42	SAGAR	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
43	SHIMLA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
44	SURAT	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
45	SURYAPET	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
46	THIRUCHIRAPALLI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
47	TIRUPATHI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
48	VARANASI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
49	VIJAYWADA	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01
50	VISAKHAPATNAM	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
51	WARANGAL	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

District Level Estimates of Agricultural Household Income in India Using Small Area Models

B. B. Singh [1]

Abstract

The paper fills in the gaps of granular data on income of agricultural households meant for capturing the disparity in income, across districts in India. NSSO has recently conducted Situation Assessment Survey of Agricultural Households in India with reference period of July 2018 to June 2019. Direct survey estimates have less precision owing to the small sample size. The paper uses small area linear mixed models to provide reliable estimates of agricultural household income by strengthening the NSSO direct survey estimates with supplementary data on population, livestock, and agriculture yield. The results show significant improvement in district estimates. Optimum combination of supplementary data has been used for different States and in some cases a group of smaller States/UTs have been clubbed together to effectively use the supplementary data for improvement of estimates.

The paper also uses spatial models which exploit the neighborhood relation between districts to further strengthen the district estimates depending on the significant spatial autocorrelation/ reduction in Akaike Information Criterion (AIC). Benchmarking of the selected model based estimates with respect to higher level estimates have been conducted and final district estimates obtained. State wise thematic maps of monthly district income of agricultural households have been made to visualize the disparity amongst the districts. Estimates have been obtained through R Package on Small Area Estimates and the district maps through bharatmaps.gov.in/makemymap.

Key Words: Agricultural households, direct survey estimates, linear mixed models, spatial models, small area estimation, spatial autocorrelation, benchmarking.

JEL Codes: C21,C33,C52,C83

Receipt of the Final Version of Paper from Author: September 2024

Acceptance: October 2024

---

[1] Director & CEO, ISS&RF, email: drbbsingh@hotmail.com

1. Introduction

In recent years, the focus of the Government has been on the welfare of the agricultural households measured through the income they have from whatever be the sources, as distinct to the production and productivity in the agriculture and allied sector. The principal, indeed, virtually only source of data on agricultural households' income is the decennial survey of National Sample Survey Office (NSSO) in the Ministry of Statistics and Programme Implementation (MoSPI). Recently the NSSO released data on income of the agricultural households based on the Situation Assessment Survey of agricultural households conducted by them during January-December 2019 for the period July 2018-June 2019.

The survey defines and covers Agricultural Households as "A household receiving value of produce more than ₹4000/- from agricultural activities (e.g., cultivation of field crops, horticultural crops, fodder crops, plantation, animal husbandry, poultry, fishery, piggery, beekeeping, vermiculture, sericulture, etc.) and having at least one member self-employed in agriculture either in the principal status or in subsidiary status during last 365 days." Income of agricultural households comprises wages, net receipts from crop production, net receipts from farming of animals and net receipt from non-farm household businesses.

NSSO does not bring and has not brought out the estimates of income below state level owing to limitation of the sample size and consequently lower precision of the estimates. A lot of research has been carried out on the level, trend, and determinants of agricultural households income at State level based on the NSS data as an aid to policies and programs. However, the implementation of the policies, Government programmes and welfare measures require estimates at district and lower levels of governance. District level estimates can be generated through the unit household level data as available from NSSO, but their reliability is a detriment to the estimates to be used for policy planning and research. Earlier, Inferential Survey Statistics and Research Foundation (ISS&RF) (2022) brought out the survey estimates at the level of NSS regions comprising districts having similar characteristics in a State. The regional estimates have higher precision and are fit for regional planning. Here, the efforts have been made to provide district level estimates of agricultural households income as more precise and reliable.

Sustainable Development Goals (SDG) emphasizes the need for granular data meant for showing how disparities within each country differs over time and to address this concern, the 2023 SDG pledged that "no one will be left behind" and called for more granular data by measuring SDG indicators for various clusters and population. Data granularity for survey based estimates implies that the Survey sufficiently represents sample for each sub-group of population. It advocated Small Area Estimation (SAE) methodology for such cases which can provide more reliable granular level estimates by borrowing strength from other data collection vehicles with more comprehensive coverage, thus artificially increasing the sample size. Agricultural households share 54% of the rural households and their welfare and disparity of income among them directly concern two Goals viz. SDG-1 "End poverty in all Its forms everywhere"; and SDG-2 "End hunger, achieve food security and improved nutrition, and promote sustainable agriculture" and the present exercise works towards the strategy and provides granular data to help towards monitoring the achievements of sustainable development targets.

2. Database

The National Sample Survey Office in its 77th round has adopted a stratified two stage sampling for the Survey, the First Stage Units (FSU) being villages/Sub-Units (SU) and the Ultimate Stage Units (USU) as households. Survey covers whole of Indian Union except villages in Andaman & Nicobar Islands which are difficult to access. The survey was conducted during January to December 2019 with reference period of Agriculture Year July 2018 to June 2019. Each sample FSU and sample Households were visited twice during the round. The first visit continued till the end of August 2019 and the second from September 2019 to December 2019. Survey was designed to have estimates for all India and State level. The sample design was made to divide each of the States into stratum and substratum (within stratum). Districts in the State constituted stratum and in each of the stratum, 3 substrata were formed and 2 FSUs from each of the stratum were drawn through Simple Random Sampling without Replacement (SRSWoR). In the selected FSU, five Second Stage Strata (SSS) were formed and 2 households from each SSS were taken. NSS Surveys have two samples, Central Sample where survey is conducted by NSSO and State Sample where Survey is conducted by the State Governments. Here the Estimates are based on Central Sample data. Sample size planned consists of 5950 Rural FSUs. However, only 5940 FSUs and 45715 households could be surveyed in visit-1 and 5894 FSUs and 44770 households in visit-2. NSSO has provided estimates for 28 States (including northeastern States) and for a group of seven north eastern States combinedly and a group of Union Territories together. State wise survey estimates of Average Monthly Income of Agricultural households and their Relative Standard Errors (RSEs) along with number of FSUs surveyed in visit-2, as per the NSS Report (2021) have been shown in Table-1.

Table-1

State wise Survey Estimates of Household Income and their Relative Standard Errors

Sr. No.	State/Union Territory	No of FSUs Surveyed in Visit-2	Household Income (₹)	RSE (%)
1	Andhra Pradesh	244	10291	9.0
2	Arunachal Pradesh	72	19225	10.3
3	Assam	231	10639	5.7
4	Bihar	518	7460	5.7
5	Chhattisgarh	122	9626	12.3
6	Gujarat	218	12578	5.8
7	Haryana	104	22220	5.2
8	Himachal Pradesh	67	12081	12.3
9	Jammu & Kashmir	41	18626	17.4
10	Jharkhand	152	4871	6.4
11	Karnataka	253	13337	6.8
12	Kerala	152	17765	5.9
13	Madhya Pradesh	335	8285	5.5
14	Maharashtra	441	11458	4.7
15	Manipur	112	11202	6.8
16	Meghalaya	78	29242	31.6
17	Mizoram	52	17912	13.7
18	Nagaland	48	9875	17.9
19	Odisha	258	5083	4.5
20	Punjab	122	24049	7.8
21	Rajasthan	336	12443	5.2
22	Sikkim	48	12444	7.8
23	Tamil Nadu	298	11852	5.8
24	Telangana	130	9336	5.9
25	Tripura	118	9894	6.4
26	Uttar Pradesh	787	7942	3.1
27	Uttarakhand	68	13361	8.6
28	West Bengal	420	6668	3.5
Group of North-Eastern States		528	18445	14.7
Group of Union Territories		51	16915	16.9
All India		5894$	10084	1.4

$ Includes Delhi and Goa States which respectively had 8 and 10 FSUs surveyed.

Source: Ministry of Statistics and Programme Implementation, National Statistical Office, Government of India. (2021). Situation Assessment of Agricultural Households and Land and Holdings of Households in Rural India 2019. NSS 77th Round (January-December 2019).

3. Survey Estimates

We have the response variable as household income per agricultural household which is the ratio of two aggregates household income Y and number of households X for the two characteristics y and x respectively. Unit level NSS data includes multiplier computed for each of the households and estimate of household income at district, State and for any group of States can be obtained just by summing up the household income multiplied by the multiplier over all the households across the strata belonging to the district/State/ group of States. In the same way, the estimates of households can be obtained by summing up of multiplier across the strata. In the notations used by NSS, estimates of household income for j[th] second stage stratum of a stratum × sub-stratum, can be expressed as $\widehat{Y}_j = \frac{N}{n_j}\sum_{i=1}^{n_j}\frac{H_{ij}}{h_{ij}}\sum_{k=1}^{h_{ij}} y_{ijk}$, j=1,…,5, estimate for a stratum sub-stratum as $\widehat{Y}_{st} = \sum_j \widehat{Y}_j$ and that for stratum as $\widehat{Y}_s = \sum_t \widehat{Y}_{st}$ and for district/State as $\widehat{Y} = \sum_s \widehat{Y}_s$.

Here subscripts s has been used for stratum, t for sub-stratum, i for FSU, j for second stage stratum in an FSU, k for sample household within an FSU, N as total number of FSUs in any sub-stratum, n as number of sample FSUs surveyed including 'uninhabited' and 'zero cases' but excluding casualty for a particular sub-stratum, H, as total number of households listed in a second-stage stratum of an FSU and h, as number of households surveyed in a second-stage stratum of an FSU. Under the above symbols, y_{stijk} is observed value of the characteristic y for the k[th] household of the j[th] second stage stratum of the i[th] FSU for the t[th] sub-stratum of s[th] stratum.

Combined ratio estimate, $\hat{R}$ of ratio $R = Y/X$ may be shown as $\hat{R} = \hat{Y}/\hat{X}$

Sampling fraction being quite small, the variance estimates using the Simple Random Sampling with Replacement (SRSWR) in place of the sampling strategy of SRSWoR has been assumed. In this case variance estimates become simple in form and there is not much loss in accuracy of variance estimates, if SRSWR is assumed. As such, estimates of Survey Variance of aggregate household income $\hat{Y}$ can be obtained as

$$\widehat{Var}(\hat{Y}) = \sum_s \widehat{Var}(\hat{Y}_s) = \sum_s\sum_t \widehat{Var}(\hat{Y}_{st}) \text{ where } \widehat{Var}(\hat{Y}_{st}) = \frac{1}{n_{st}(n_{st}-1)}\sum_{i=1}^{n_{st}}\left(N_{st}\hat{Y}_{sti} - \hat{Y}_{st}\right)^2$$

And that of income per agricultural household ratio $\hat{R} = \frac{\hat{Y}}{\hat{X}}$ as

$$\widehat{MSE}(\hat{R}) = \frac{1}{\hat{X}^2}\sum_s\sum_t \widehat{MSE}_{st}(\hat{R})$$

$$\text{where } \widehat{MSE}_{st}(\hat{R}) = \frac{1}{n_{st}(n_{st}-1)}\sum_{i=1}^{n_{st}}\left[N_{st}\left(\hat{Y}_{sti} - \hat{R}\hat{X}_{sti}\right) - \left(\hat{Y}_{st} - \hat{R}\hat{X}_{st}\right)\right]^2$$

$N_{st}\hat{Y}_{sti} = \sum_s\sum_t y_{stijk} \times n_{st} \times multiplier$ and $N_{st}\hat{X}_{sti} = \sum_s\sum_t x_{stijk} \times n_{st} \times multiplier$. Multiplier at stratum/sub-stratum/second stage stratum level has been given as $\frac{N_{st}}{n_{stj}} \times \frac{H_{stij}}{h_{stij}}$. $j = 1,2,3,4,5$. Estimates of Relative Standard Error (RSE) may be shown as $\widehat{RSE}(\hat{R}) = \frac{\sqrt{\widehat{MSE}(\hat{R})}}{\hat{R}} \times 100$

The formulae of the estimated MSE at district/State depends on the estimates at the stratum X sub-stratum level. There are few sub-strata which have single FSU. In this case we cannot compute MSE and therefore the MSE has been computed based on only those sub-strata which have more than one FSU. Similarly, there are a few districts wherein none of the stratum have more than one FSU. In these cases, MSE is based on the variance of households in the related FSUs/stratum X sub-stratum. Details can be seen as a Note on Sampling Design and Estimation Procedure, contained in Appendix C of the MoSPI (2021).

4. Small Area Models and Estimates

Small area models and techniques to strengthen the small area survey estimates rest on the exploitation of relationship with neighboring and similar small areas as evident through auxiliary information available for the small areas. Here, a simple area level linear mixed model, Fay and Herriot model (1979) has been employed which strengthens the estimates through common regression parameters. It has been referred to as Linear Mixed Model here in this paper.

$$y = \theta + \varepsilon, \quad E(\varepsilon|\theta) = 0, \quad Var(\varepsilon|\theta) = R$$

$$\theta = X\beta + v$$

$$y = X\beta + v + \varepsilon \tag{1}$$

where θ is a m-component vector (corresponding to number of small areas, here districts) for household income per household and y is direct survey estimator, obtained through small sample surveyed data, $X(m \times p)$ is a design matrix of auxiliary variables, $\beta(p \times 1)$ is column vector of regression parameters including intercepts, and ε and v are respectively sampling errors and random effects assumed to be independently distributed. For the estimation of parameters, it has been assumed as $v \sim N_m(0, \sigma^2 I)$ and $\varepsilon \sim N_m(0, R)$. R is a diagonal matrix of order m with elements as survey estimates of Mean Squared Errors (MSE) for the mth small area/ district.

In Indian context, it is difficult to find out auxiliary variables having higher correlation coefficient with the response variable, in this case district household income as more than 0.5%. It limits the improvement in the estimates. Therefore, spatial relationship between the districts in the form of neighborhood relations has been exploited on the idea that neighboring districts share similar social and economic characteristics. The model is referred to as Spatial Model. It takes the form

$$y = \theta + \varepsilon, \quad E(\varepsilon|\theta) = 0, \quad Var(\varepsilon|\theta) = R$$

$$\theta = X\beta + u \quad\quad u = \rho W u + v \quad\quad E(v) = 0, \;\; Var(v) = \sigma^2 I$$

$$y = X\beta + Zv + \varepsilon \quad\quad Z = (I - \rho W)^{-1} \tag{2}$$

$W(m \times m)$ is a known spatial weight matrix and shows the amount of interaction between any pair of small areas. The constant $|\rho| < 1$ is a measure of overall level of spatial autocorrelation and its magnitude reflects suitability of W for given y and X. The parameter vector $\varphi = [\rho, \sigma^2]^T$ has two elements. Matrix W has been formed in such a way that, $W_{ij} = 1$ (unscaled) $\forall i, j = 1,2 \dots m$ if jth area is physically contiguous to ith area and 0 otherwise. $W_{ii} = 0 \;\forall i = 1,2 \dots m$. The matrix has been standardized as to satisfy $\sum_{j=1}^{m} W_{ij} = 1 \;\forall i = 1,2 \dots m$.

Best Linear Unbiased Predictor (BLUP) of the true value of household income $y = X\beta + Zv$ and the MSE of the BLUP have been worked out by using linear mixed model approach Prasada and Rao (1990), Datta and Lahiri (2000), Singh, Shukla and Kundu (2005).

$$\hat{\beta} = [X^T\Sigma^{-1}X]^{-1}X^T\Sigma^{-1}y \quad\quad \hat{v} = \sigma^2 Z^T \Sigma^{-1}\left[y - X\hat{\beta}\right]$$

$$\Sigma = \sigma^2 A + R \quad A = (I - \rho W)^T (I - \rho W)$$

$$\hat{\theta}(\psi) = X\hat{\beta}(\psi) + \Lambda(\psi)\left[y - X\hat{\beta}(\psi)\right] = \sigma^2 \mathrm{A}^{-1}(\psi)\Sigma^{-1}(\psi)y + R(\psi)\Sigma^{-1}(\psi)X\hat{\beta}(\psi)$$

$$= g_1(\psi) + g_2(\psi)$$

$$\Lambda(\psi) = \sigma^2 \mathrm{A}^{-1}(\psi)\Sigma^{-1}(\psi) = R(\psi)\Sigma^{-1}(\psi)X\hat{\beta}(\psi)$$

BLUP estimator depends on the parameters ψ which is not known. It has been estimated from data through Restricted maximum Likelihood Estimation. Substitution of ψ by its estimator $\hat{\psi}$ makes Empirical Best Linear Unbiased Predictor (EBLUP) as

$$\hat{\theta}(\hat{\psi}) = g_1(\hat{\psi}) + g_2(\hat{\psi})$$

and the second order approximation to the MSE of the EBLUP can be obtained as

$$MSE[\hat{\theta}(\hat{\psi})] = g_1(\hat{\psi}) + g_2(\hat{\psi}) + g_3(\hat{\psi}) + o(m^{-1}) \tag{3}$$

$$g_1(\hat{\psi}) = \Lambda(\psi)\, R = R - R\Sigma^{-1}R$$

$$g_2(\hat{\psi}) = R\Sigma^{-1}X[X^T\Sigma^{-1}X]^{-1}X^T\Sigma^{-1}R$$

$$g_3(\hat{\psi}) = \sum_{d=1}^{q}\sum_{e=1}^{q} L_d\,(\psi)\Sigma(\psi)L_e^T(\psi)I_{de}^n(\psi)$$

$\mathrm{L}(\psi) = \underset{1 \le d \le q}{Col}[L_d(\psi)]$, $L_d(\psi) = \frac{\partial\Lambda(\psi)}{\partial\psi_d}, d = 1,2,\ldots,q$ and $I_\psi^{-1} \equiv \left(I_{de}^n(\psi)\right)$ for d,e=1,2…,q

and Estimator of the MSE as

$$mse[\hat{\theta}(\hat{\psi})] = g_1(\hat{\psi}) + g_2(\hat{\psi}) + 2g_3(\hat{\psi}) - g_4(\hat{\psi}) - g_5(\hat{\psi}) + o(m^{-1}) \tag{4}$$

where $E\left[mse\left(\hat{\theta}(\hat{\psi})\right)\right] = MSE[\hat{\theta}(\hat{\psi})] + o(m^{-1})$

$$g_4(\hat{\psi}) = \frac{1}{2}\sum_{d=1}^{q}\sum_{e=1}^{q} I_{de}^n(\psi)\, Trace\left[I_\beta^{-1}(\psi)\frac{\partial I_\beta(\psi)}{\partial\psi_d}\right]\frac{\partial g_1(\psi)}{\partial\psi_e}$$

$$g_5(\hat{\psi}) = \frac{1}{2}\sum_{d=1}^{q}\sum_{e=1}^{q}\left[R\Sigma^{-1}(\psi)\frac{\partial^2\Sigma(\psi)}{\partial\psi_d\partial\psi_e}\Sigma^{-1}(\psi)RI_{de}^n(\psi)\right]$$

$\frac{\partial g_1(\psi)}{\partial\psi}$ is a partition matrix of order $(mq \times m)$ having q matrices of order $(m \times m)$. In the same way $\frac{\partial^2\Sigma(\psi)}{\partial\psi\psi^T}$ is a partition matrix of order $(mq \times mq)$ having q partitions row and column wise with $\frac{\partial^2\Sigma(\psi)}{\partial\psi_d\partial\psi_e}$ as a general submatrix of order $(m \times m)$ therein. For any squared partitioned matrix $B = \underset{1 \le d \le q}{Col}\left[\underset{1 \le e \le q}{Concat}(B_{de})\right]$ with square sub-matrices of same order we have $\underset{m}{Trace}(B) = \sum_{d=1}^{q} B_{dd}$

The expression gives the matrix of estimator of the MSE of EBLUP, $\hat{\theta}(\hat{\psi})$ and the MSE of the individual small area estimator may be obtained as the respective diagonal elements in the matrix. The terms $g_3(\hat{\psi})$, $g_4(\hat{\psi})$ and $g_5(\hat{\psi})$ are the contributions, due to the estimation of parameter vector ψ by $\hat{\psi}$. In case of simple model without spatial autocorrelation, the term $g_5(\hat{\psi})$ becomes zero.

5. Response Variable, Auxiliary Variables, Diagnostic Tools, and Tests for Selection of Models

In India, the administrative data suffers from coverage, accuracy and uniform availability at district level. Population Census, conducted decennially provides district level data, related to the response variable uniformly with sufficient details for all the States/UTs. Similar is the case for livestock census. Most recent census data available pertains to the year 2011, however, it works as better covariates than many other variables related to response variable. A host of auxiliary variables from Census of India 2011 related to literacy and work force, and agriculture yield for foodgrains 2018-19 and animal husbandry per thousand households from 19th Livestock Census 2012 have been identified to regress the district average household income (survey income). Results of 20th Livestock Census are yet to be released. Agricultural household income comes from four different components, through cultivation, animal farm, wages and non-farm business which make the modelling of response variables with the available auxiliary information more challenging. From the list of 30 variables and odd, a combination of auxiliary variables was selected by step wise regression method resulting with the lowest Akaike Information Criterion (AIC) and substantial significant reduction in AIC. Step wise regression largely prevents multicollinearity problem. In the process, F statistics with related p value, adjusted multiple correlation coefficient R^2, significance of β coefficients and diagnosis of residuals through $q - q$ plot have been studied, and the most suitable group of auxiliary variables decided. Care has been taken to judiciously include the predictors whose relationship with the response variable can be explained and which have better correlation with the response variable. Besides a tradeoff between goodness of model fit and model complexity with more predictors has been tried and not more than two predictors have been considered however without losing goodness of fit. In a few cases, the step wise regression could not find any exogenous variable and invariably selects the intercepts only. In such cases further insights of the working force at the disaggregated 2-dgit NIC level has been attempted and better exogenous variables selected. Regression exercise to find out the most suitable set of exogenous variables have been carried out for individual States/UTs.

In the State of Uttar Pradesh, one district Ballia has been found with rare survey estimate of monthly income per household as just ₹504.90 (State income per household ₹7942) with RSE of the estimate as 1201.27%. Abysmally low income is due to the component receipt of non-farm business as (-) ₹5559.06. Further analysis finds that one household in an FSU for Visit-1 had non-farm expanses of ₹1.20 lakh and receipt of ₹1000. Further, the household had a multiplier of 20382. This household non-farm business net receipt was modified to the average net receipt of other non-farm business (total 5 in visit 1). This exercise modified the survey estimate of income for Ballia as ₹6939.92 and the RSE of the estimate as 10.33%. This modification has been made to the database and no other changes have been made. The exercise has slightly changed the State estimate of income too.

As far as neighborhood relation amongst the districts are concerned, a neighborhood matrix having 0,1 (1 for the neighboring districts and 0 otherwise) has been formed and scaled in such a way that sum of row total is one. The linear mixed model has been fitted on the set up of selected exogenous variables and the spatial model on the same set of selected variables along with neighborhood matrix separately and improved district estimates with variance estimates obtained.

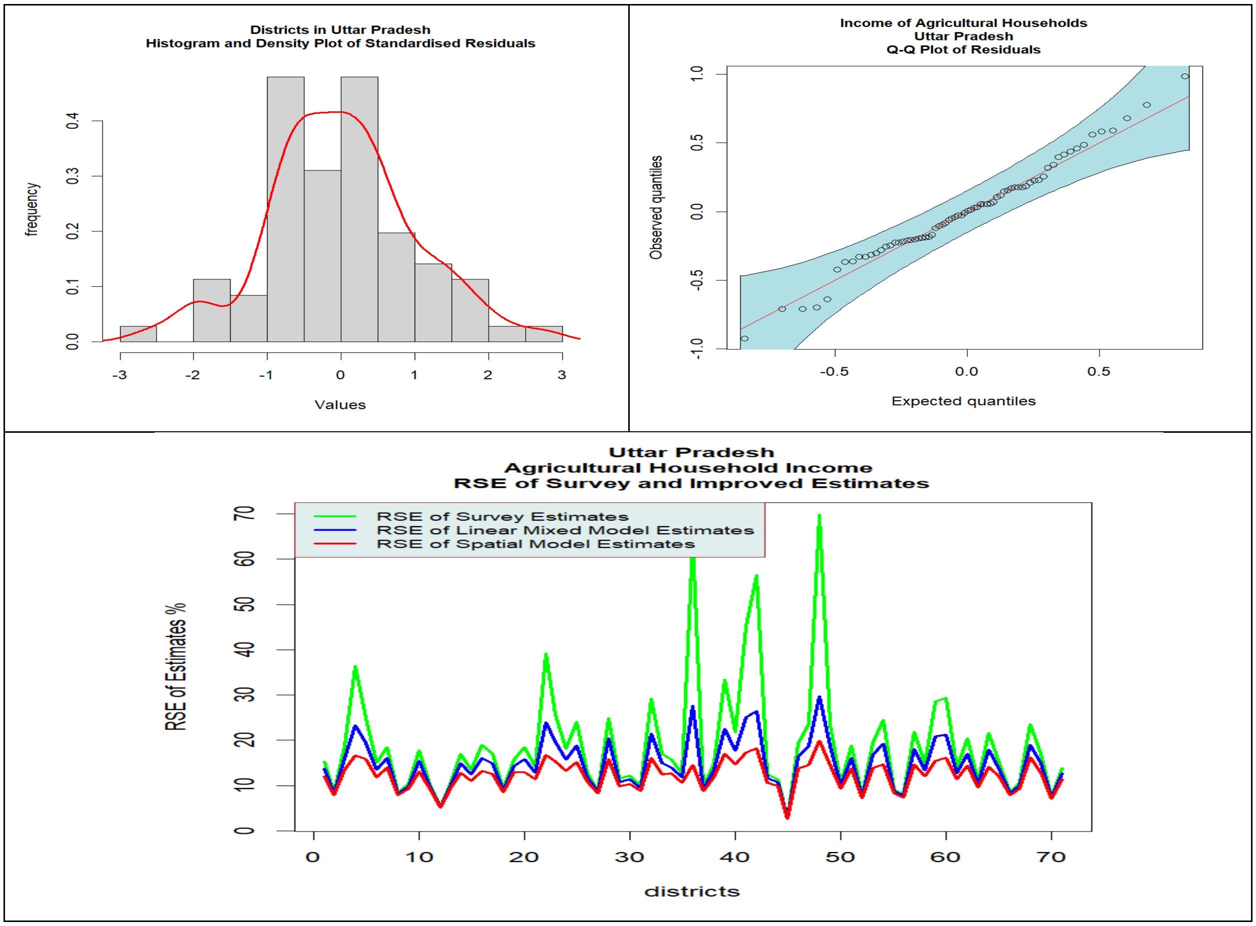
Survey household income, logarithm of the survey household income and corresponding survey based MSE with respect to auxiliary variables and logarithm of auxiliary variables have been modelled and normality of residuals tested through Shipro-Wilk Normality test. It has been found that logarithm of the survey household income works well. Adjusted R^2 for different group of States/UTs have been found in the low range of 0.20 to 0.45, questioning the availability and selection of appropriate set of auxiliary variables. However, given the databases largely available with statistical system in India, auxiliary variables have been studied and the most appropriate variables were selected for the models. And based on them, the F statistics and p value of the linear regression coefficients and those of linear mixed and spatial models have been found which are generally highly significant for all the group of States/UTs. Results of diagnosis for each group of States/UTs along with histogram, density plots of standardized residuals, Q-Q plot of residuals and RSE of the survey and model based estimates are available with the author and at the ISS&RF site https://issrf.in. As an example, for Uttar Pradesh, the same has been presented in Table 2 and figure 1.

Table 2
Diagnosis Results of Modelling for Uttar Pradesh

Exogenous Variables: % Rural Main Workers engaged in Plantation, Livestock, Forestry, Fishing, Hunting and allied activities 2011 and Crop Yield Foodgrains 2018-19				
F Statistics	DF	p value	Multiple R2	Adjusted R2
17.47	2,68	7.53E-07	0.3441	0.3248
	Estimate	Std. Error	t Value	Pr(>|t|)
(Intercept)	8.0751	0.24836	32.513	<2.00E-16
X12	0.29448	0.06016	4.895	6.34E-06
X20	0.14916	0.08269	1.804	0.0757
	beta	std.error	tvalue	pvalue
	Linear Mixed Model			
(Intercept)	7.8819374	0.24338334	32.384869	4.48E-230
X12	0.2778703	0.05500199	5.052004	4.37E-07
X20	0.2175474	0.08165644	2.664179	7.72E-03
	Spatial Model			
(Intercept)	8.3125319	0.27781487	29.921119	1.05E-196
X12	0.1949088	0.05693758	3.423201	6.19E-04
X20	0.1212791	0.08308037	1.45978	1.44E-01

Goodness	Loglike	AIC	BIC	Spatial Corr
Linear Mixed	-23.92694	55.85389	64.9046	
Spatial	-13.89064	37.78128	49.09468	0.8245539
Shipro-Wilk Normality Test			W=0.98561	P value=0.5946

Figure 1
Histogram and Density Plot of Standardised Residuals, Q-Q Plot of Residuals and RSE of Survey and Improved Estimates of Agricultural Household Income of Districts in Uttar Pradesh



The model used here is

$$log(y) = X\beta + Zv + \varepsilon \quad Z = (I - \rho W)^{-1} \text{ with } (\varepsilon|\theta) = 0 \;\; and\; Var(\varepsilon|\theta) = R/y^2 \qquad (5)$$

Thus, we have used the transformation in equations (1) and equation (2) as $y \rightarrow log(y)$ and $R \rightarrow R/y^2$ and after estimation of the transformed household income and its MSE, back transformation has been used to find out the estimator and its variance as

$$\hat{\theta} \rightarrow exp(\hat{\theta} + mse(\hat{\theta})/2) \;\text{ and }\; mse(\hat{\theta}) \rightarrow exp(mse(\hat{\theta}) - 1) \times exp\left(2(\hat{\theta} + mse(\hat{\theta})\right)$$

The estimates of the parameters and the mean squared error (MSE) of the estimates have been obtained through R Package on Small Area Estimates as developed by Isabel Molina and Yolanda

Marhuenda [2015] with slight changes. The package is based on the research papers by Singh, B. B. et all [2005], Prasad, N. N., and J.N.K. Rao. [1990] and Datta, G. S. and Lahiri, P. [2000].

Likelihood Ratio Test (LRT) $-2\log L \sim \chi_1^2$ has been used for comparison of Linear Mixed and Spatial Model. LRT is the ratio of likelihoods at the hypothesized parameter values for two competing models under different hypothesizes $H_0: \rho = 0$ Vs $H_1: \rho \neq 0$. Besides Akaike Information Criterion (AIC) and Bessel Information Criterion (BIC) have been utilized for selection of models. It may be understood that Spatial Model uses an additional parameter ρ which is estimated from the data and therefore it may increase the Mean Standard Errors (MSE) resulting estimates of some of the districts are less reliable than those obtained through linear mixed models. Ultimately the significance of parameters, distribution of errors terms and resultant average Relative Standard Errors (RSE) of the fitted model are studied and the improved estimates obtained.

6. Grouping of States and Union Territories

Sizeable number of States such as Andhra Pradesh, Jammu & Kashmir, Himachal Pradesh, Goa, Uttarakhand, North-Eastern States and Union Territories have smaller number of districts and for a few the diagnostic tools fail and for other few the model for estimation of MSE do not converge. In such cases, based on the proximity of States/UTs and similar distribution of agricultural household income, a group of States have been combined, and the models are fitted on the combined number of districts. This takes benefit of neighborhood relation of the adjoining districts of combining States/UTs. Smaller States/ UTs have been combined with the adjoining larger States and North-Eastern States have been separately combined as two groups. The combination of States/UTs has followed exercises carried out to model fitting separately for the individual States/UTs and the diagnostic results thereon. As such 17 Groups of States and UTs have been formed. Group wise RSE of the estimates has been shown in Table 3.

Table 3
Group of States/UTs, their Composition and Average RSE of Survey, Linear Mixed and Spatial Model Estimates and the Selected Models Based on Log Likelihood and AIC

Sr. No.	Group of States and UTs	No of Districts	Average RSE of Estimates (%)			Selected Model
			Survey	Linear Mixed	Spatial	
1	Andhra Pradesh and Telangana	43	30.08	15.73	16.03	Linear Mixed
2	Assam	27	24.09	14.40	14.90	Linear Mixed
3	Bihar and Jharkhand	62	26.84	17.27	17.24	Linear Mixed
4	Chhattisgarh	27	45.69	22.31	23.75	Linear Mixed
5	Gujarat, Dadra & Nagar Haveli, and Daman & Diu	36	29.56	17.73	17.96	Linear Mixed
6	Haryana, Punjab, Chandigarh, and Delhi	46	22.76	15.62	15.82	Linear Mixed
7	Jammu & Kashmir, Himachal Pradesh, and Uttarakhand	35	56.87	25.25	23.65	Spatial
8	Karnataka, Goa, Lakshdweep Islands and Kerala	47	31.34	19.79	20.85	Linear Mixed
9	Madhya Pradesh	50	39.31	18.99	19.11	Linear Mixed
10	Maharashtra	33	26.95	16.48	15.35	Linear Mixed
11	Manipur, Nagaland, Tripura, and Sikkim	28	34.14	19.29	20.10	Linear Mixed
12	Arunachal Pradesh, Mizoram, and Meghalaya	31	62.14	51.56	55.14	Linear Mixed
13	Odisha	30	22.48	14.11	11.69	Spatial
14	Rajasthan	33	32.13	18.37	18.83	Linear Mixed
15	Tamil Nadu, Puducherry, and A&N Islands	36	24.19	20.11	19.40	Spatial
16	Uttar Pradesh	71	27.27	16.69	13.85	Spatial
17	West Bengal	22	17.18	11.16	11.21	Linear Mixed
All India		657	-	-	-	-

Linear Mixed and Spatial Models both have shown improvement in district estimates of agricultural households. Further Spatial Model, depending on the strength of spatial autocorrelation, wherever fitted well based on LRT and AIC, provides better estimates in comparison to that of Linear Mixed Model. Table 4 presents frequency distribution of districts State wise by the level of RSE of Survey based estimates and out of such districts with RSE exceeding 20%, the number of districts having RSE less than 20% as per Linear Mixed and Spatial Model. Result for all India shows that 37.36% districts with Survey based estimates with RSE exceeding 20% have now RSE as less than 20% under Linear Mixed Model. On the other hand, 42.86% districts with Survey based estimates with RSE exceeding 20% have now RSE as less than 20% under Spatial Model.

Table 4

Frequency distribution of districts by the level of RSE of Estimated Average Household Income

Sr. No.	State/UT	Total no. of districts	No. of districts by the level of RSE of Survey based estimate				Out of such districts with RSE exceeding 20%, the no. of districts having RSE < 20% as per	
			10% or less	10% - 20%	20% - 30%	30% or more	Linear Mixed Model	Spatial Model
1	Andaman & Nicobar Islands	3	-	2	-	1	-	-
2	Andhra Pradesh	13	2	8	1	2	1	1
3	Arunachal Pradesh	16	4	3	3	6	-	-
4	Assam	27	6	9	9	3	12	12
5	Bihar	38	10	16	8	4	4	4
6	Chandigarh	1	1	-	-	-	-	-
7	Chhattisgarh	27	8	10	3	6	-	-
8	Dadra & Nagar Haveli	1	-	-	1	-	1	1
9	Daman & Diu	2	1	-	-	1	-	-
10	Delhi	2	2	-	-	-	-	-
11	Goa	2	-	1	-	1	-	-
12	Gujarat	33	6	10	11	6	8	8
13	Haryana	21	5	9	5	2	4	4
14	Himachal Pradesh	12	3	3	3	3	-	-
15	Jammu & Kashmir	10	-	2	3	5	-	-
16	Jharkhand	24	3	10	7	4	3	3
17	Karnataka	30	3	12	5	10	3	3
18	Kerala	14	3	7	2	2	2	2
19	Lakshdweep Islands	1	-	-	-	1	-	-
20	Madhya Pradesh	50	11	14	9	16	9	9
21	Maharashtra	33	6	10	11	6	14	16
22	Manipur	9	3	3	2	1	-	-
23	Meghalaya	7	1	4	-	2	-	-
24	Mizoram	8	1	-	4	3	-	-
25	Nagaland	11	2	2	1	6	1	1
26	Odisha	30	6	13	7	4	9	11
27	Puducherry	2	-	-	1	1	-	-
28	Punjab	22	1	12	3	6	2	2
29	Rajasthan	33	2	16	7	8	6	6
30	Sikkim	4	1	2	1	-	-	1
31	Tamil Nadu	31	5	14	7	5	3	3
32	Telangana	30	9	9	5	7	4	4
33	Tripura	4	1	2	1	-	1	1
34	Uttarakhand	13	5	1	4	3	3	3
35	Uttar Pradesh	71	13	36	15	7	11	21
36	West Bengal	22	9	11	1	1	1	1
All India		657	133	251	140	133	102	117

7. Bench Marking of Estimates

State wise survey estimates of household income have been published by the MoSPI and it is in public domain. Relative Standard Errors (RSE) of these estimates (at State level) are also low resulting in reliable estimates. There is a need for coherence between aggregated model-based estimates (district level) and the higher (State) level direct survey estimates. As income per household at district level cannot be aggregated, we have used district estimates of income and district estimates of household for the benchmarking of estimates. Estimates of households have been kept unchanged and it has been assumed that the model only changes income estimates. Ratio adjustment approach has been adopted for benchmarking of the Income estimates at district level.

Benchmark Income per Household for a district (d) in the State has been obtained as

$$B_d = \left[\ M_d \times \frac{\sum_d Y_d}{\sum_d M_d \times X_d}\right] \tag{6}$$

where M_d represents model-based estimates of income per household, X_d represents survey-based estimate of number of households and Y_d survey based estimates of aggregate household income for a specific district d. Summation of d is over all the districts in the State. Summation of X_d and Y_d over all the districts in the State is the survey estimates of aggregate household income and aggregate households in the State, their ratio representing estimates of income per household at State level.

Benchmarking of district level estimates have been computed through linear mixed or spatial models depending on the results for test of model selection. The benchmark estimates are the final estimates of the districts. Benchmarking of the district household incomes for major States (excluding north eastern States) have been worked out based on the respective State household income while district income in north eastern States (Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura) are based on combined north eastern States income and the districts income of Union Territories (Andaman & Nicobar Islands, Chandigarh, Dadra & Nagar Haveli, Daman & Diu, Lakshdweep Islands and Puducherry) based on combined income of all the Union Territories together. NSSO provides household income for these two groups of States/UTs separately.

8. District wise Estimates of Household Income of the Agricultural Households and District Maps

State level Tables containing district wise Survey Estimates, Estimates based on Linear Mixed and Spatial Models with their Relative Standard Errors (RSE) and Benchmark and Final Estimates of Agricultural Household Income district wise is available with the author and at the ISS&RF site https://issrf.in. The Tables have average State estimates based on district survey estimates and estimates based on models with their RSEs. Benchmark and Final Estimate at State level are the survey estimates based on large samples. There may be slight differences in the benchmark and final estimates at State level with that of survey based RSE shown in Table 1. As an example, for Uttar Pradesh, district wise agricultural household income has been presented in Table 5.

Table 5
Uttar Pradesh
Average Monthly Income of Agriculture Households
District Wise Survey and Small Area Estimates and their Relative Standard Errors (RSE)

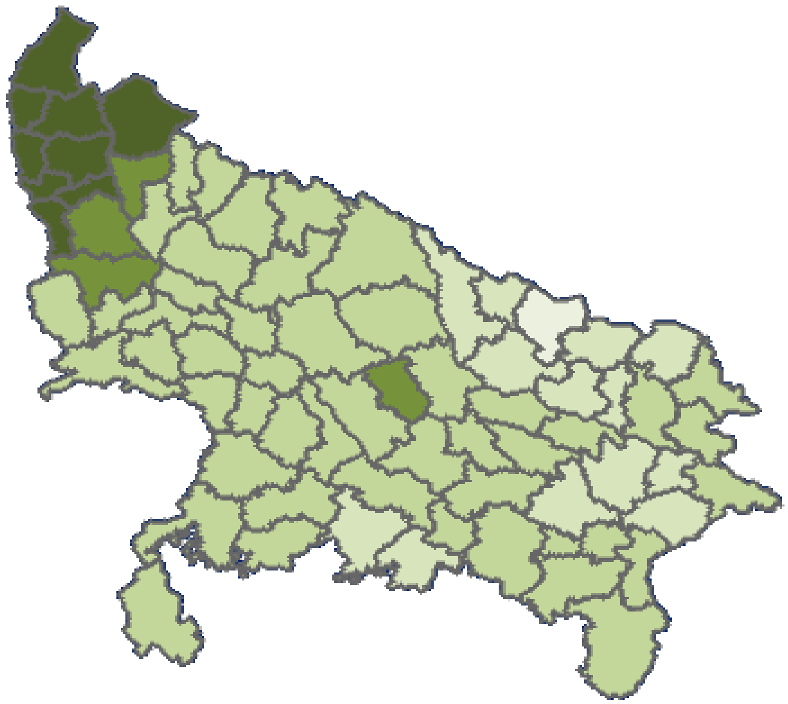
(Selected Model is Linear Mixed and Benchmarked and Final Estimates are based on Linear Mixed)

Sr.	District	Income Estimates (₹)			RSE of Estimates (%)			Benchmark and Final Estimates (₹)
		Survey	Linear Mixed	Spatial	Survey	Linear Mixed	Spatial	
1	Agra	6554.54	7056.56	6985.55	15.27	13.69	12.06	7015.05
2	Aligarh	10368.16	10576.35	10538.61	8.47	8.21	7.79	10583.11
3	Allahabad	7549.50	7798.58	8310.45	18.64	15.95	13.51	8345.54
4	Ambedkar Nagar	7554.93	8003.32	6721.83	36.51	23.33	16.67	6750.22
5	Auraiya	8873.74	8174.52	7318.65	24.94	19.37	15.83	7349.56
6	Azamgarh	6526.16	6714.24	6058.38	15.08	13.55	11.81	6083.96
7	Baghpat	15726.58	14303.04	15798.74	18.47	16.06	13.99	15865.45
8	Bahraich	4615.31	4694.61	4871.58	8.52	8.22	7.79	4892.16
9	Ballia	6939.92	7248.07	6806.97	10.33	9.83	9.28	6835.72
10	Balrampur	2981.25	3645.87	3906.87	17.82	15.45	13.08	3923.37
11	Banda	5393.34	5386.84	5649.07	10.25	9.77	9.31	5672.92
12	Bara Banki	7387.95	7365.84	7322.98	5.30	5.22	5.12	7353.90
13	Bareilly	7533.90	7534.36	7520.83	10.17	9.67	9.17	7552.59
14	Basti	6660.11	6748.76	6065.00	16.98	14.89	12.88	6090.61
15	Bijnor	21271.70	18600.32	18345.01	13.54	12.42	11.03	18422.47
16	Budaun	6463.60	6622.15	7086.11	18.96	16.14	13.20	7116.04
17	Bulandshahr	12246.27	12328.27	13273.98	17.00	14.99	12.53	13330.03
18	Chandauli	8116.13	8229.12	8015.53	9.27	8.89	8.47	8049.38
19	Chitrakoot	6154.46	5770.79	5858.15	15.75	14.30	13.02	5882.88
20	Deoria	7151.08	7231.94	6225.34	18.53	15.90	12.99	6251.62
21	Etah	7119.93	7159.94	7018.00	14.12	12.86	11.44	7047.63
22	Etawah	3227.51	6171.52	7008.37	39.22	23.97	16.80	7037.96
23	Faizabad	6201.56	6903.57	6807.77	25.68	19.63	15.32	6836.52
24	Farrukhabad	7342.23	7421.11	7386.23	18.03	15.67	13.19	7417.42
25	Fatehpur	5559.71	6429.97	6984.94	24.11	18.91	15.13	7014.44
26	Firozabad	8815.22	8605.98	7963.83	12.82	11.85	10.81	7997.46
27	Gautam Buddha Nagar	21025.56	20851.44	20154.49	8.83	8.59	8.23	20239.60
28	Ghaziabad	21503.86	21347.08	23561.33	24.86	20.42	15.89	23660.82
29	Ghazipur	5458.04	5664.07	5424.27	11.47	10.76	9.76	5447.18
30	Gonda	4926.94	5086.86	4927.15	12.24	11.41	10.27	4947.96
31	Gorakhpur	6606.13	6802.56	6455.65	9.89	9.43	8.77	6482.91
32	Hamirpur	11572.30	8457.70	7403.45	29.19	21.52	16.15	7434.72
33	Hardoi	7057.51	6860.64	7233.26	16.99	14.91	12.50	7263.80
34	Jalaun	5288.58	5573.84	6346.22	15.73	14.02	12.71	6373.01
35	Jaunpur	5405.76	5625.39	5845.39	12.72	11.77	10.54	5870.07
36	Jhansi	5762.69	6298.40	8414.79	64.41	27.68	14.52	8450.32
37	Jyotiba Phule Nagar	14520.37	13647.90	13542.17	9.54	9.12	8.71	13599.36
38	Kannauj	6727.03	6814.01	6775.87	14.92	13.48	11.88	6804.48
39	Kanpur Dehat	12164.01	9566.66	7799.10	33.48	22.54	17.05	7832.04
40	Kanpur Nagar	4399.38	5619.98	6719.19	21.76	17.74	14.64	6747.56
41	Kanshiram Nagar	6317.23	8141.32	7941.29	45.05	25.01	17.28	7974.82
42	Kaushambi	18496.33	8615.72	7969.78	56.47	26.45	18.19	8003.43

Sr.	District	Income Estimates (₹)			RSE of Estimates (%)			Benchmark and Final Estimates (₹)
		Survey	Linear Mixed	Spatial	Survey	Linear Mixed	Spatial	
43	Kheri	7970.74	7867.01	7829.02	12.44	11.59	10.56	7862.07
44	Kushinagar	6741.62	6815.77	6568.41	11.38	10.68	9.99	6596.14
45	Lalitpur	8111.88	8079.71	8087.72	2.59	2.58	2.58	8121.88
46	Lucknow	12500.88	11142.45	10637.38	19.26	16.33	13.74	10682.30
47	Mahamaya Nagar	6565.28	7718.93	7578.11	23.57	18.70	14.51	7610.11
48	Mahoba	11267.00	6440.48	7375.96	69.89	29.80	19.95	7407.10
49	Mahrajganj	5515.27	6466.13	6001.63	23.58	18.69	14.55	6026.97
50	Mainpuri	7130.43	7155.02	7107.51	10.39	9.90	9.30	7137.53
51	Mathura	6756.08	7559.34	7266.95	18.82	16.07	13.70	7297.64
52	Mau	3706.90	3905.46	4066.43	7.82	7.59	7.25	4083.60
53	Meerut	19041.66	18477.83	21922.20	19.34	16.76	13.92	22014.76
54	Mirzapur	11056.85	9580.20	8530.37	24.52	19.21	14.64	8566.39
55	Moradabad	9125.40	9115.42	9641.72	9.16	8.79	8.36	9682.43
56	Muzaffarnagar	17546.70	16945.20	17294.14	8.09	7.85	7.31	17367.17
57	Pilibhit	7668.84	7449.45	7585.03	21.82	18.03	14.63	7617.06
58	Pratapgarh	8149.71	8004.91	7784.74	14.77	13.33	11.99	7817.61
59	Rae Bareli	4363.28	5437.08	6720.55	28.59	20.86	15.47	6748.93
60	Rampur	5500.49	6236.15	6691.87	29.40	21.29	16.08	6720.13
61	Saharanpur	13468.15	12598.92	14007.49	13.83	12.64	11.43	14066.64
62	Sant Kabir Nagar	3290.96	4209.00	4639.90	20.50	17.07	14.37	4659.50
63	Sant Ravidas Nagar (Bhadohi)	6525.85	6770.55	6932.18	10.98	10.35	9.52	6961.45
64	Shahjahanpur	7739.90	7436.47	7466.30	21.59	17.91	14.17	7497.83
65	Shrawasti	4725.44	4967.03	4558.64	14.93	13.47	11.87	4577.89
66	Siddharthnagar	5894.64	6024.57	5720.70	8.49	8.21	7.90	5744.85
67	Sitapur	7965.60	7691.38	7657.66	10.60	10.04	9.31	7689.99
68	Sonbhadra	5019.85	5371.00	6106.41	23.60	18.97	16.25	6132.20
69	Sultanpur	7669.60	7895.25	7593.52	16.98	14.90	12.84	7625.59
70	Unnao	7645.16	7550.75	7458.76	7.49	7.29	7.06	7490.25
71	Varanasi	9472.28	9187.66	8635.99	13.86	12.66	11.36	8672.45
Uttar Pradesh Average		8446.52	8250.68	8335.63	27.27	16.69	13.85	8041.28

District thematic maps have been created for each of the State through bharatmaps.gov.in of the Ministry of Electronics and Information Technology. Thematic maps use State tables and have distribution of household income in five different ranges of Less than Mean-3SD/2; Mean-3SD/2 to Mean-SD/2, Mean-SD/2 to Mean+SD/2, Mean+SD/2 to Mean+3SD/2, greater than Mean+3SD/2 to show the distinct deviations in the district income. The Map shows 785 district boundaries. In contrast NSSO has only 657 districts based on the 2011 census. Also a few districts viz. Central, East, New Delhi, North, North East, Shahdara, South, West, South East (Delhi), Mirpur, Muzaffarabad, Anantnag, Barmulla, Badgam, Bandipura, Ganderbal, Kulgam, Kupwara and Shopian (Jammu & Kashmir), Leh Ladakh and Kargil (Ladakh), Hyderabad (Telangana), Mumbai and Mumbai Suburban (Maharashtra), Chennai (Tamil Nadu), Kolkata and 24 South Paraganas (West Bengal) where survey could not be conducted due to them being entirely urban areas or otherwise have been shown as not available. Thematic maps for each of the State is available with the author and at ISS&RF site https://issrf.in. One of such maps for Uttar Pradesh has been presented as figure 2.

Figure 2
District Wise Distribution of Monthly Income in Uttar Pradesh



<=₹2100	₹2100 - ₹6100	₹6100 - ₹10000	₹10000 - ₹14000	>₹14000	Not Available

There are a sizeable number of States wherein few districts have been created/ bi-trifurcated/ merged with other districts after the Census 2011. In these cases, composition of then 2011 districts were ascertained through the Government notifications and new districts in the map have been shown with the same household income as that of the 2011 districts. In a few cases where new districts comprise of parts of more than one 2011 districts, average income of the part districts has been shown. Still there are few districts in the National Sample Survey that do not correspond with the 2011 census, the typical districts being from Telangana carved out from Andhra Pradesh after the 2011 census but before the NSS Survey. In such cases district census handbooks of the erstwhile districts have been referred to and exogenous variables computed. Models discussed here have the potential to provide estimates for the new districts created after Census 2011 without the survey estimates with the help of exogenous variables available for the new districts. However, due to the non-availability of census data for the newly created districts we could not estimate average income for them.

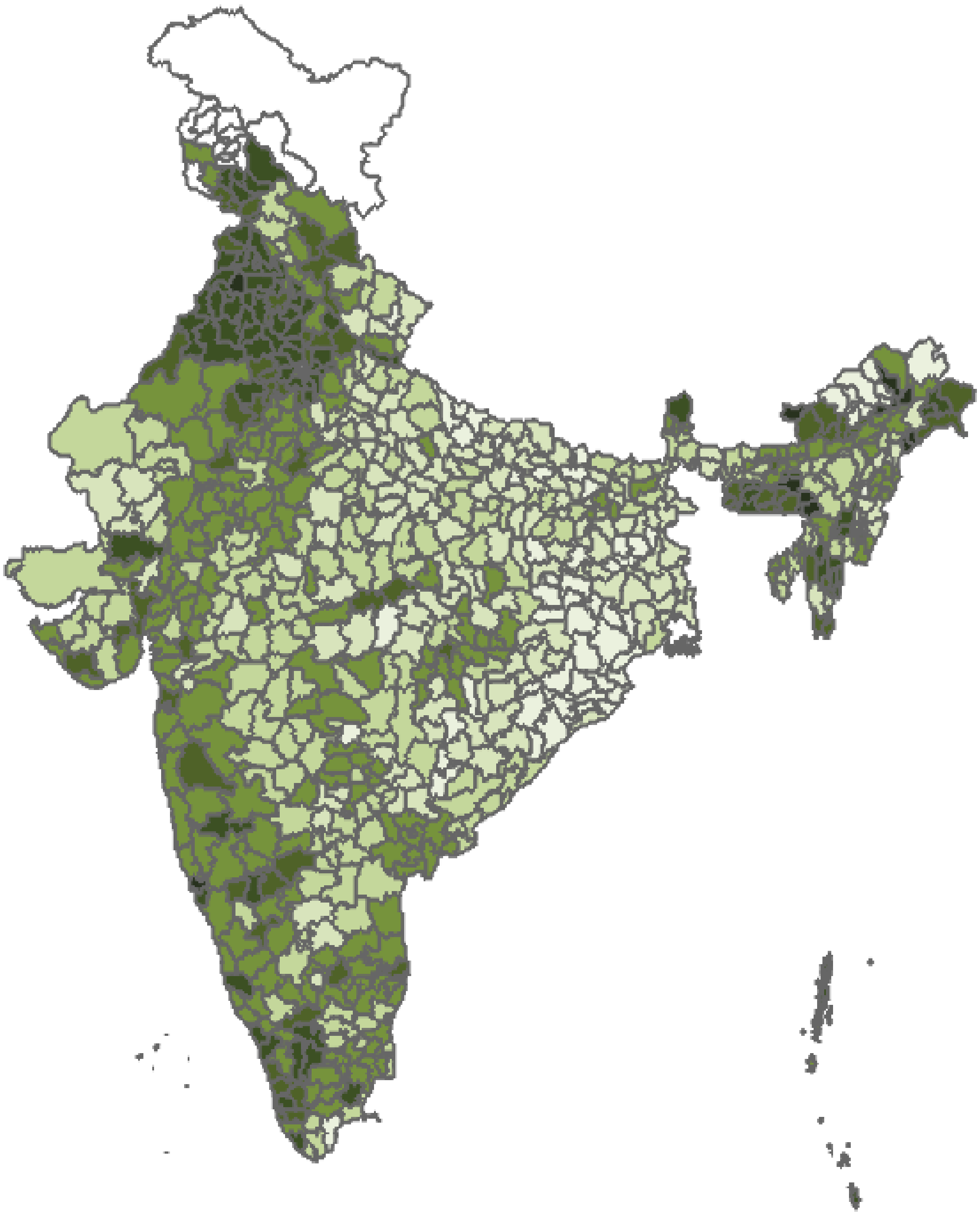
State wise frequency of districts in different ranges of household income has been shown in Table 6 and an All-India Map with district wise distribution of household income for agricultural households, in Figure 3.

Table 6
State Wise frequency of Districts in Different Household Income Ranges (₹)

	<5000	5000-7500	7500-10000	10000-15000	15000-20000	20000-30000	>=30000	Total
Andaman & Nicobar Islands	-	-	-	-	1	1	1	3
Andhra Pradesh	-	1	7	5	-	-	-	13
Arunachal Pradesh	5	-	-	2	5	1	3	16
Assam	-	1	13	11	1	1	-	27
Bihar	4	16	14	4	-	-	-	38
Chandigarh	-	-	-	-	-	-	1	1
Chhattisgarh	4	7	6	9	1	-	-	27
Dadra & Nagar Haveli	-	-	-	1	-	-	-	1
Daman & Diu	-	-	-	1	-	1	-	2
Delhi	-	1	-	1	-	-	-	2
Goa	-	-	1	-	-	1	1	2
Gujarat	-	2	7	18	5	1	-	33
Haryana	-	-	-	-	4	17	-	21
Himachal Pradesh	-	-	4	5	3	-	-	12
Jammu & Kashmir	-	-	-	2	3	5	-	10
Jharkhand	15	9	-	-	-	-	-	24
Karnataka	-	1	4	17	6	2	-	30
Kerala	-	-	-	5	5	4	-	14
Lakshadweep	-	-	-	-	-	1	-	1
Madhya Pradesh	1	14	28	5	2	-	-	50
Maharashtra	-	-	17	14	1	1	-	33
Manipur	-	3	1	3	1	1	-	9
Meghalaya	-	-	-	-	3	2	2	7
Mizoram	-	1	-	-	4	3	-	8
Nagaland	1	3	2	5	-	-	-	11
Odisha	10	18	2	-	-	-	-	30
Puducherry	-	-	1	1	-	-	-	2
Punjab	-	-	-	1	2	18	1	22
Rajasthan	-	2	6	19	3	3	-	33
Sikkim	-	-	-	3	-	1	-	4
Tamil Nadu	1	-	9	16	1	4	-	31
Telangana	1	7	12	10	-	-	-	30
Tripura	-	-	2	2	-	-	-	4
Uttarakhand	1	3	4	2	1	2	-	13
Uttar Pradesh	6	36	18	5	3	3	-	71
West Bengal	1	16	5	-	-	-	-	22
All India	50	141	163	167	55	73	9	657
	7.61	21.46	24.81	25.42	8.37	11.11	1.37	100.00

Figure 3

All India District Wise Distribution of Monthly Income of Agricultural Households



<₹5000	₹5000 - ₹7500	<=₹7500 -₹10000	₹10000 - ₹15000	₹15000 - ₹20000	₹20000 - ₹30000	>₹30000	Not Available

9. Conclusions

District Survey Estimates, Estimates based on Linear Mixed and Spatial Models, their Relative Standard Errors (RSE) and the Benchmarked Estimates of Household Income for Agricultural Households have been made and Benchmarked Estimates have been worked out on the Estimates selected, based on AIC and LRT. The following conclusions are drawn.

i) There is dearth of auxiliary variables related to agricultural household income. However, it has been found that Population Census could provide adequate number of exogenous variables having multiple correlation more than 0.25 which can be modelled to provide improved estimates.

ii) Logarithmic transformation of response variable household income has better correlation with auxiliary variables and the models based on transformed variable has resulted residuals having normal distribution.

iii) States require different optimum sets of auxiliary variables, different measures of neighborhood relation amongst the districts and different grouping of States/UTs together to enhance the reliability of estimates.

iv) Two competing models Linear Mixed and Spatial have been used to find improved district estimates and one out of the two has been selected based on likelihood ratio test and AIC. It was possible to reduce the RSE significantly and enhance the reliability of estimates for most of the districts by using linear mixed or spatial models. States where spatial autocorrelation has been found significant or AIC get lowered in comparison to the linear mixed model, the spatial model further reduces the RSE of the estimates.

v) Spatial Models have an additional neighbourhood relation which has been exploited, however, due to estimation of additional parameter of autocorrelation, in a few cases, RSE of the estimates get slightly higher to that of linear mixed models.

vi) It has been found that direct survey estimates of 58.45% districts have less than 20% RSE which increases to 73.36% districts for linear mixed and to 75.19% for spatial one. Linear Mixed and Spatial Models respectively have significant number of districts 26.8% and 24.8% with RSE more than 20% and 4.25% districts with RSE more than 30% in each case. Estimates of such districts still lack reliability; however reliability is a relative concept and to that extent, the models provide significant gain for the districts having survey estimates with more than 20% or 30% RSE. Spatial models have been found better than Linear Mixed for 4 States/ Group of States viz., Jammu & Kashmir, Himachal Pradesh, and Uttarakhand; Odisha, Tamil Nadu, Puducherry, and Anaman & Nicobar Islands; and Uttar Pradesh. Frequency distribution of districts by RSE of the estimates under different models have been shown in Table 7.

Table 7
Frequency Distribution of Districts by RSE of Household Income Estimates

Survey and Models	RSE of District Estimates						% Share
	<10%	10-20%	20-30%	30-50%	>=50%	Total	<30%
Survey	133	253	140	98	35	657	79.76
Linear Mixed	144	337	154	19	3	657	96.65
Spatial	150	344	141	18	4	657	96.65

vii) Linear Mixed and Spatial Models based district estimates have uniformly improved the survey estimates, still they have high RSEs as summarised in Table 2 and therefore one needs to read the district figures with caution along with their respective RSEs.

viii) There are 9 groups of States/UTs with contiguous boundaries and with similar distribution of household income for which unified model have been used. The models reveal better fit than for the States individually. North-Eastern States have been categorised in two groups depending on the outcome of regression analysis corroborated with final model outcomes.

ix) Analysis of Benchmarked Estimates of State wise household income has revealed that 53.73% districts have average income less than Rs.10,000 while only 12.48% districts have more than Rs.20,000 per month. These 12.48% districts are mostly in the Northeastern States/ Union Territories and in Haryana and Punjab.

x) The National Statistical Commission and the Statistics and Economic Ministries should encourage for generation of improved estimates of important indicators as part of their activities and tools for evidence-based governance and decision making. This also requires effective monitoring of Sustainable Development Goals.

Acknowledgements

Unit level data for the paper was made available by Ministry of Statistics and Programme Implementation, and the base district maps by National Informatics Centre, Government of India. The collaborative project on “Analysis of Structure and Determinants of farmers Income” conducted jointly by ISS&RF and ICRIER motivated the paper. The paper acknowledges the contribution of the Team ISS&RF to whom the paper was referred to and discussed by them threadbare from conceptualization to the final product, and to the referees for providing suggestions to improve.

References:

[1] Asian Development Bank (2020). Introduction to Small Area Estimation Techniques, A Practical Guide for National Statistics Offices. https://cran.r-project.org/web/packages/sae/sae.pdf

[2] Bell, W. R., Datta, G. S. and Ghosh, M. (2013). Benchmarking Small Area Estimators. Biometrika, 100, 1, 189-202

[3] Chandra, Hukum. (2013). Exploring Spatial Dependence in Area-Level Random Effect Model for Disaggregated-Level Crop Yield Estimation. Journal of Applied Statistics. Volume 40, Issue 4, 823-842. https://doi.org/10.1080/02664763.2012.756858.

[4] Datta, G. S. and Lahiri, P. (2000). A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. Statistica Sinica. 10, 613-627.

[5] Datta, G.S., Ghosh, M., Steorts, R. and Mapler, J. (2011). Bayesian Benchmarking with Applications to Small Area Estimates. Springer, 20, 574-588.

[6] Eurostat. (2019). Guidelines on Small Area Estimation for City Statistics and other Functional Geographies.

[7] Fay, R. and R. Herriot. (1979). Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. Journal of the American Statistical Association. 74(366a), 269-277.

[8] Government of India, Ministry of Agriculture, Department of Animal Husbandry, Dairying and Fisheries. (2014). 19th Livestock Census-2012. All India Report.

[9] Government of India, Ministry of Agriculture and Farmers Welfare, Directorate of Economics and Statistics. Area, Production and Yield (API) 2018-19. https://data.desagri.gov.in/website/crops-apy-report-web.

[10] Government of India, Ministry of Home Affairs, Office of the Registrar and Census Commissioner of India. Census of India 2011. https://censusofindia.gov.in/census.website/data/census-tables.

[11] Government of India, Ministry of Statistics and Programme Implementation, National Statistical Office. (2021). Situation Assessment of Agricultural Households and Land and Holdings of Households in Rural India 2019. NSS 77th Round (January-December 2019).

[12] Government of India, Ministry of Statistics and Programme Implementation, National Statistical Office. Unit Level data & Report on NSS 77th Round for Schedule- 33.1, January 2019 – December 2019, (Land and Livestock Holding of Households and Situation Assessment of Agricultural Households). https://mospi.gov.in/unit-level-data-report-nss-77-th-round-schedule-331- january -2019 -%E2%80%93-december-2019land-and-livestock

[13] Hidiroglou, Michael. (2017). Small Area Estimation: Theory and Practice. Statistics Canada, Statistical Innovations and Research Division. 3445-3456.

[14] Inferential Survey Statistics and Research Foundation. (2022). Analysis of Structure and Determinants of Farmers Income. A Project conducted under the Joint Research Agreement between Inferential Survey Statistics and Research Foundation and Indian Council for Research on International Economic Relations.

[15] Molina, Isabel, and Marhuenda, Yolanda. (2015). SAE: An R package for small area estimation. The R Journal. 7(1), 81-98.

[16] Pereira, Luis N. and Coelho, Predro S. (2012). Small Area Estimation using a Spatio-Temporal Linear Mixed Model, REVSTAT-Statistical Journal, Volume 10, Number 3, 285-308.

[17] Prasad, N. N. and J.N.K. Rao. (1990). The Estimation of the Mean Squared Error of Small Area Estimators. Journal of the American Statistical Association, 85(409), 163-171.

[18] Rao, J.N.K. and Molina, I. (2015). Small Area Estimation. Wiley Series in Survey Methodology. John. Wiley and Sons Inc. Hoboken, New Jersey.

[19] Singh, B. B., Shukla, G. K. and Kundu, D. (2005). Spatio-Temporal Models in Small Area Estimation. Survey Methodology 31, 183-195.

[20] United Nations Population Fund. (2020). Better Data for More Effective Policies and Programmes: Small Area Estimation.

[21] Young, Linda J. and Chen, Lu. (2022) Using Small Area Estimates to Produce Official Statistics. Stats, 5(3),881-897. https://doi.org/10.3390/Stats503005.

Females in the Job Market: An Understanding of Its Socio-Economic Correlates

Atanu Sengupta[1] Priyanka Das[2] and Shenaz Hussain[3]

Abstract:

Women's employment is an important aspect of empowerment. To understand female labour force participation, it is important first to analyse the factors that affect women's participation in the labour market. For the present study, we have used PLFS (Periodic Labour Force Survey) data in 2022-23 at the unit level. We have used a multinomial logistic model for our purpose. Our analysis reveals the importance of socio-economic factors in shaping women's employment. The entire analysis is based on rural and urban differences. We have considered the variables region, religion, social group, head of household, household type, educational level, age, and marital status of women. These affect both rural and urban women's employment in different dimensions of magnitude.

Key words: women worker, rural-urban, region, socio-economic factors, PLFS unit level data at 2022-23

JEL Code: J01, J08, J16, J24

Receipt of Final Version of paper from Author: November 2024

Acceptance: November 2024

[1] Professor, Department of Economics, The University of Burdwan, Email: asengupta@buruniv.ac.in
[2] Research Scholar, Department of Economics, The University of Burdwan, Email: das1994priyanka@gmail.com
[3] Research Scholar, Department of Economics, The University of Burdwan, Email: shenazhussain89@gmail.com.

1. Introduction:

One of the important issues of the development paradigm is the role of gender empowerment (Sen, 1984, 1987, 1993). This issue has its importance from two perspectives: ethical perspectives and development perspectives. It is argued by many thinkers that ethically it would be wrong to depress almost half of the population deprived and in duress (Plato, Republic V; Mills, 1869; Wollstonecraft, 1792). Such discrimination has no logical basis.[4] Hence there is a sufficient background to look at the essential lack of ethical metre in dealing with women's issues.

For social thinkers, however, the development is more persuasive. Depriving women in human capital formation is depriving the possibility of growth and development that can precede such kind of development. This is not beneficial to the development process. There is an interesting study by Chattopadhyay and Duflo (2004) based in India, the study focuses on the impact of political reservation of women in the local bodies. The study shows that the development of women's leadership in the village council makes an important impact on the infrastructure formation at the village level. This is because when women become decision-makers they make decisions regarding the improvement of rural infrastructures. The influence of women decision-makers on village infrastructure is largely positive. Their focus on initiatives that benefit the entire community supports sustainable and inclusive development, making women’s leadership a valuable asset for rural progress. Empowering women in leadership positions, therefore, contributes significantly to improving quality of life and building stronger, more vibrant communities. In contrast, when males become decision-makers, they emphasise other issues that give more personal benefits.

Women's employment is an important aspect of empowerment. There are many ways in which women's employment is a crucial factor towards empowering women. First, it generates an income that is capable of paying many needs of her own. In this sense, it gives her a type of freedom. Second, women's employment uses the work effort of women in a more socially

[4] ‘But then, I said, as we have determined to speak our minds, we must not fear the jests of the wits which will be directed against this sort of innovation; how they will talk of women's attainments both in music and gymnastic and above all about their wearing armour and riding upon horseback!’ Plato, *Republic*, Book V

desirable way than otherwise possible. Third, it endorses a sense of dignity to the women than it otherwise achievable (Sen, 2001).

Among all the components of empowerment, social thinkers give the stress on employment. It is argued that women's position in society and their empowerment usually depend on their significant role in economic activity. The anthropological literature by Friedrich Engels (2010) has long argued that in primitive societies women play a crucial role in livelihood and economic decisions that guarantee they are omnipotent in such societies. However, with the growth of private property and the family, women's works were relegated. Despite social significance, the privatised household care activities did not gain much social recognition. Women were degraded.

In ancient times women played a significant role among the lower strata of the population compared to the upper class (Habib, 2007). With the waves of the Industrial Revolution, a larger amount of women tend to participate in the labour market. This happened at all works from minimal to top intellectual. There has been argued by human development theory (Amartya Sen 2001) that such employability raises their empowerment and brings about a change in their dignity[5]

There are many attempts to study women's empowerment in various types of ways and usages. It is to be noted that the arrival of women in the labour market is meted through several factors. The problem is very complex for a country like India with wide socio-economic variation. Such differences also prop up differently across the length and breadth of the vast country. In this study, we have tried to plot out such differences as they are manifested across several socio-economic variates. The present study attempts to analyse the various factors which affect women's employment in India. These factors wish to cover the niceties of the Indian panorama.

To understand female labour force participation, it is important first to analyse the factors that affect women's participation in the labour market. Women's labour supply is crucially affected by the fact of time constraints because women all over the world tend to spend more time in

[5] Some feminist thinkers (Maria Mies 1982) have argued that meagre job market inclusion does not improve women's empowerment. She studied the lace makers of Nasrapur in Andhra Pradesh. Many of these lacemakers are housewives. They work as “part-time workers” within the ambit of a housewife. Hence they do not get the dignity of a labourer. Sengupta and Mukherjee (2010) found in their study that women labourers in the bidi industry are not getting any identity as ‘labour’ though they play a significant labouring activity. This is because bidi binding means supposed to carry “leisure activity” or “part-time activity” among household chores. However, they argue that the welfare of these women labourers can be improved only if they are given the recognition of workers and not removing them from the labour market. Employment in the labour market is a necessary condition for women's empowerment though not sufficient. It acts in a significant way towards empowering her position in the family and the states

unpaid domestic chores and caregiving activity, which fall outside of the standard definition of economic activity by labour statistics (Hirway, 2012).

Typically in India, men are regarded as the main breadwinners of the family **(Kabeer,N.(2012), Nanda,P., Datta, and Das,P.(2014))**. While men's participation in the job market depends mainly on likelihood concerns, women's participation in the labour market is influenced by various demographic, reproductive, social, religious and cultural factors. The Present study intends to carry out a detailed analysis of various factors such as employment, unemployment and those not in the labour force based on various dimensions such as age, education, HH type household type, caste, religion, marital status, and family income etc, using PLFS data from 2022-23.

The paper is organized into six sections. In the next section, we have analysed a brief literature review. The third section analyses the data and methodology of our study. The fourth section examines the activity status of women on various dimensions by descriptive analysis. In the fifth section, we analyse our regression results. In the last section, we conclude our results and suggest some policies to improve their status.

2. A Brief Review of Literature:

Employment is an important issue. There are many facets of it. Various researchers concentrated on different aspects of these issues. It is not possible to cover all those things in a single paper. We will concentrate only on those issues that are relevant to the present study. In the present study, we are dealing with the individual decision of offering oneself for employment and the factors that affect it. Though entering into the labour market may be an individual decision, it is often constrained by social, economic, familial, and cultural factors. Many of these factors are not identifiable but still, we can select at least a set of such factors which may be important in explaining the spread of women's employment. In the literature review, we concentrate on the role of these factors, studied by various researchers.

Women's participation in the labour market is an essential component of gender equality. It can create different developmental objectives including educational attainment, health care accessibility, reducing fertility, maternal mortality and child mortality; increasing nutritional levels, , and increasing average age at marriage(Dreze and Sen,2013; Costagliola,2021).

In India, at present, the U-shaped hypothesis of female LFPR and economic growth is not occurring because the re-entering of females in the labour market is not happening **(**Sarkar et al., 2018; Sundari, 2020; Costagliola,2021)**.** Costagliola (2021) argues that this outcome is a result of traditional and male-dominated perspectives on the societal role of women. He also believes that the transition from agriculture to the services sector does not happen successfully because it is difficult for them to balance domestic responsibility and job-related tasks. He also opines that female employees in the particularly informal sector are in danger of experiencing sexual abuse due to the failure to enforce the 2013 sexual harassment rights of women in the workplace. For this reason, many women feel unsafe to participate in the labour markets. Also, there are various **demand side** and **supply side** factors which explain that reason. We now deal with some papers which concentrate on these issues. Women's **labour demand** is constrained by human capital possession and discriminatory effects. Bergman's **(**1974) overcrowding model shows that two separate occupations for males and females have arisen due to employers' preference for discrimination against women, which leads to differences in their job opportunities and these job opportunities are lower for females. Besides this Sundari (2020) argues that employers intentionally differentiate the 'male job' and 'female job 'and fix the lower wage for female job to evade the law of equal remuneration act,1976 according to which it is mandatory to pay equal wages for equal work.

Investment in education and health varies for boys and girls due to differing future expectations for each gender. Traditionally it is expected that boys will become 'bread earners', so investment in such human capital accumulation of boys is essential, while girls seem to be future 'homemakers' rather than paid workers in the labour market (PROBE, 1999; Tilak et al. 2002; Rustagi, 2005). So human capital is low for women.

Next, we consider the **supply side factor** of women in the labour market. Women's labour supply is also constrained by different supply-side factors. The following factors are pointed out in the study of Mehrotra & Sinha (2017). a) **Increase** in household income: Female labour market participation influenced by income effect. As the HH income increases, the necessity of women to remain labour market reduces, and they withdraw from the labour market. b) Mechanization in agriculture: In the agriculture sector women perform more manual work than men. The processes of mechanization in the agriculture sector reduce the demand for female labour. c) Increase in educational enrolment: There is an increase in girls' enrolment for the age group below 15 as well as 15-19 yrs after 2005 which has reduced women's participation in the labour market since 15 is the legal age for working. d) Increase in domestic responsibility: domestic duties, care caregiving activity are constraints to women to participate in the labour market. Since the enrolment in secondary school increased, the task of younger sibling care performed by the mother as well as an increase in the nuclearisation of the family suffers from a lack of support from other family members who are constrained to join the women's labour market.

Swaminathan (2020) argues that there is also some measurement bias that low female WPR, as reported by labour force survey, may be misleading. she argues that standard labour force surveys underestimate women’s work based on various perspectives including the nature of women’s work which is home-based, remittent, and in the informal sector and she suggests that by expanding the definition of ‘work’ including the women who reported themselves as ‘attending domestic duties’ then the tendency of low labour participation of women may be reduced some extend. Hirway (2012) argues that with the help of time-use surveys, a large part of women's labour is missing standard NSSO data which does not mean the withdrawal of women from the labour market.

Along with these, there are various **demographic and social factors** which influence women's participation in the labour market. Mehrotra and Parida (2021) show in their study that age is a proxy of work experience, so it has a positive influence on labour market participation. They also point out that married and unmarried women are less likely to participate in the labour market due to social stigma and patriarchal norms. They also view that the nature of jobs which women do are quite different. Women belonging to economically backward classes including poor, SC, and ST normally work in the agriculture, construction, labour labour-intensive manufacturing sectors but not working is a matter of prestige for better-off households and other upward classes. Considering the religions of women study shows that Hindu and Muslim women are less likely to participate in the labour market compared to others including Christian and Sikh women. Studies also exhibit that LF participation of women is higher in southern and western states in India due socio socio-economic developments and cultural norms as compared to women in central and eastern regions where women have less freedom to move out for jobs outside their locality and still these regions are agrarian based. Reed (2020) views that more generally in rural areas, unmarried women are less likely to participate in the labour market than married women because of social restrictions and security concerns but in urban areas, the situation is the opposite because too many women like to invest their time in household (HH) chores. Interestingly he notes that rural married women are now following the behaviour of urban women as HH income has increased, so employment has fallen in rural areas among the married women.

Education plays an important role in participating women in the labour market. The relationship between educational attainment and women's work participation is very complex. There is positive and negative and both types of relationships are shown. The study by Andres et al (2017) shows a U-shaped hypothesis between women's education and their labour market participation. They point out that, for lower levels of education, women's participation in the labour market increases but it declines with increasing educational level, after that trend it increases with higher level of education. They explain this. At very low levels of education, women are compelled to join the work by necessity condition if their household income is very low. At the middle level of education, they face restriction by social and patriarchal constraints as well they are reluctant to join low-skill work. At higher levels of education, women are free from cultural norms. Brinton, Lee and Parish (1995) opine that education is related to employment through both income and substitution effects. Educated women marry educated men, who have higher incomes. Since high-income families have a lesser need for women's contribution to the household, this encourages women to withdraw from the labour force. On the other hand, educated women also have higher incomes than less educated women, which encourages them to participate in the labour force.

Dhanaraj & Mahambare (2017) point out that education enhances women's decision-making power. They can decide to join the labour market. They also point out that Family pressure against work is reduced for highly educated women due to higher earning capacity and quality of job.

Ghai (2018) points out some phenomena that relate to higher education associated with low labour market participation---a) Higher quality of education improves the marriage prospects of females rather than enhance employability prospects. Many parents and girls see education as value to improve marriageability condition. b) High patriarchy index associated with low lab market participation for higher educated women. c) Women remain in the outside labour force to continue their education. Reed (2020) also points out that a working wife is undesirable to many Indians while an educated wife is desirable.

In brief, we have considered the demand side, supply side, demographic side and educational side of women employment. These are very important antenna that captures the wavelength of our paper. We try to answer many of these questions faced by the researchers in this divert but related arena. Before going to the analysis, we move to data and methodology in the next section.

3. Data and Methodology:

3.1. Data:

For the present study, we have used PLFS (Periodic Labour Force Survey) data in 2022-23 at the unit level. It is designed to estimate the key employment and unemployment indicators. This is contrary to the low-skill quinquennial survey which gave us employment data only after a certain period. However, it has been criticised for its frequency. PLFS help us to provide the data with almost yearly in a more frequent way (annually and quarterly) (Papola,2014). This facilitates the research in the arena of employment and unemployment including women. During the data collection of PLFS, the NSSO has introduced computer-assisted personal interviews where data is predominantly subject to validation during the survey process. Regarding the sampling frame, the recent UFS blocks that were accessible to FSU in urban regions were included. Around 50 per cent of towns based on the UFS frame are recognised as urban conglomerates according to Census 2011, which makes better coverage in terms of the total sample. Apart from this aspect, NSSO incorporates multiple variables in PLFS rounds coverage of data regarding skill and education has become more comprehensive, gives the data on earning of self-employment, and also provides data on hours worked and available for work, including the third gender. So we consider here for analysing the present scenario of employment, PLFS 2022-23data can be used (Jajoria and Jadav,2020).

PLFS provides the data as per usual status (principal+ subsidiary) as well as the current weekly status in both rural and urban areas annually. The survey covered the whole of Indian states and union territory except the villages of Andaman & Nicobor Islands.

It uses multistage stratified random sampling for collection of data. The first stage units (FSU) are census villages in the rural sector and urban frame survey block (UFS) for the urban sector. The ultimate stage is household (HH).

It covers the total sample size of 12,800 FSUs (7,024 villages and 5,776 UFS blocks) at all India levels.

In each hg (hamlet group)/sb (sub-block), Second-stage stratification (SSS) is done based on the number of members who have completed a secondary level of education.

The present study uses PLFS, 2022-23-unit level data on some selected variables on women's broad status (such as employment, unemployment and not in the labour force) like age, education, HH income, caste, religion, marital status, HH type, employment status etc.

3.2. Methodology:

In the present study, we wish to focus on women's employment and various factors that affect it. We also analyse women's unemployment (Searching job or being available for a job) and not in the Labour force. As already discussed in our study of literature, we want to concentrate on various demographic, social and household characteristics of a person. These include region, religion, social cast, HH (Household) size, educational level, family income, marital status etc. We want to emphasize the variation in women's employment based on these characteristics. We have considered the working age group population only (15-59 yrs). Our analysis is carried out in two parts--- descriptive part as well analytical part. In the entire paper, we have considered the place of residents (rural and urban) separately. This is necessary because the nature of jobs available in rural areas is different from those that are available in urban areas. In rural areas, most of the informal jobs are centred around agriculture and Argo-based activities. In the urban areas informal sector jobs take the form of services, retailing on certain goods (vegetables, fish, working on housemaids and so on). Because of the difference in the job, it is not rational to treat them within a homogeneous set.

Standard regression about employment and unemployment is generally made on an aggregative basis. However, to enter into the labour market is an individual choice decision. Therefore, in this paper, we have concentrated on the individual decision-making process. In the PLFS round, there are three different states of individual--- employment, unemployment and not in labour force. These three categories are so distinct that they have to be treated separately. It was Chaudhary and Verick (2014); Mitra et al. (2020)[6] who introduce a multinomial model to deal with this inherent distinction of these three states of individual. We are also using the same methodology to analyse women's employment. The Multinomial logistic regression (MLR) model is an important method when our categorical response variable consists of more than two categories. It is capable of forecasting categorical response variables using continuous and /or categorical independent variables. The model permits us to simultaneously compare multiple contrasts that is log odds of three or more distinctions are estimated at the same time.

---

[6]Mitra et al.(2020) uses the dummy variables within the Multinomial Logit Model.

In contrast to OLS regression, multinomial logistic regression does not require the assumption of linearity between dependent and independent variables, and does not assume normally distributed variables, homoscedasticity. However, it mandates that the observation remain independent and the independent variable should display a linear relationship with the logit of the dependent variable. (Habil,2012)

4. Preliminary Findings:

Activity status of women such as employed, unemployed and not in the labour force varies across regions, social groups, religions, different demographic characteristics (i.e. age, educational level, marital status) etc. Irrespective of these we know that more women tend to remain out of the labour force followed by workers, unemployed. In this study, we analyse those dimensions which affect the status of women based on PLFS data in 2022-23.

We see that irrespective of region, the proportion of women workers is higher in rural areas than in urban areas, whereas the proportion of those seeking or available for jobs is higher in urban areas compared to rural areas. Compared to different regions we observe that the proportion of women workers is greater in the Western and Southern regions in rural areas, whereas in urban areas, the contribution is higher in the Eastern, Southern, and Western regions. This result reveals that in more developed regions, women's participation as workers is greater in both rural and urban areas, with higher participation in rural areas compared to urban counterparts.

This is led by the fact that natures of job availability which absorb women is greater in rural areas than in urban areas. On the other hand, education among urban women is greater this leads to the fact that until finding a better job women remain either as unemployed or prefer to remain in the domestic domain rather than joining any type of low-paid, low-skill job. It is also observed that overall women's labour market participation is less compared to that of in the rural area. This indicates in urban areas income effect is stronger. Generally, HH income is greater in urban areas which discourages them from participating in the labour market. They prefer to perfectly bring up their children, do HH chores etc. However urban women unemployment is greater than in rural areas. This may be due to generally among urban women educational level is high so they try to find out better-skill job until they find it they prefer to remain unemployed.

Next, we consider the women's activity status across different religions of people. People in India belong to different religious groups. They have different cultures, social norms as well as different economic conditions. We hope to see that this differentiation also affects their labour market participation. Considering different religious groups of women we see that the proportion of female workers is greater among Buddhists and Jains in rural areas, Whereas in urban areas, the proportion of female workers is greater among the Buddhist and Christianity communities. It is the least among Islam's unreserved community in both areas. Another observation is that in both areas, the proportions of women who are seeking or available for jobs are greater among Sikhs in rural areas and Buddhists in urban areas compared to other religious women (Table 1). The results reveal the fact that even the unreserved category of the Muslim community faces rigid social and patriarchal restrictions so they cannot move out to join the labour market. Also, they are less educated which is required to enter into labour market. On the other hand, women from Christianity, Buddhism and other higher religions women face less restriction and patriarchal constraints. So their labour market participation increases. Along with this social spending on this Christian and Buddhist community is greater which makes them potential for labour market participation.

Comparing different castes of women, we also observe that the proportion of female workers is high among Schedule Tribe (ST) in both places of residence whereas it is least among other forward (Table 1). The results express the fact that women from the lower cast can join any type of work for their needs because they belong to the poor section. But not working is a prestigious issue for higher cast women.

Comparing different types of household (HH) it is observed that the proportion of women employment is highest among self-employed in rural areas whereas in urban areas it is from Causal labour HH households (Table 1). This reveals the fact that rural women from self-employed HH are compelled to do work due to necessary conditions, generally, they join low-skill types of jobs. This is true for Casual labour HH in urban areas.

We also analyse the fact that if women have more decision-making powers, then what happens to their participation in the labour market? For this purpose, we have considered the head of the HH and more interestingly the results show that the proportion of women workers from female self-headed HH is much higher compared to women from other HH headed in both areas (Table 1). The result expresses the fact that generally women who become head of the HH, by dire necessity participate in any type of work or because of HH needed, women from that type of HH have to join the labour market.

Table 1: Distribution of women (15-59 years.) by activity status (US (PS+SS)) according to different socio economic and demographic variables in 2022-23 (weighted)

	Rural				Urban			
	Workers	Seeking or available for job	Not in labour force	Total	Workers	Seeking or available for job	Not in labour force	Total
Region					**Region**			
EAG	43.00	1.00	56.00	**100**	34.00	2.40	63.60	**100**
Northern	44.83	1.76	53.41	**100**	36.00	2.70	61.30	**100**
Western	50.81	0.95	48.24	**100**	39.95	1.74	58.31	**100**
Southern	47.81	1.65	50.54	**100**	38.78	2.62	58.6	**100**
Eastern	45.26	1.19	53.55	**100**	40.71	2.17	57.12	**100**
North-Eastern	45.78	1.67	52.55	**100**	38.62	2.86	58.52	**100**
Religion					**Religion**			
Hindu upper caste	45.61	1.41	53	**100**	37.84	2.28	59.9	**100**
Muslim unreserved	40.97	0.89	58.16	**100**	34.35	2.37	63.3	**100**
Christianity	45.18	1.87	52.96	**100**	38.42	2.89	58.69	**100**
Sikhism	41.71	2.33	55.97	**100**	35.78	3.17	61.05	**100**
Jainism	47.16	0.82	52.04	**100**	35.99	0.61	63.4	**100**
Buddhism	52.14	1.24	46.64	**100**	38.64	4.43	56.93	**100**
Others	45.45	1.85	52.71	100	34.76	2.53	62.71	100
Social group					**Social group**			
ST	50.25	1.26	48.51	**100**	37.94	3.11	58.95	**100**
SC	43.86	1.33	54.82	**100**	39.05	2.71	58.24	**100**
OBC	44.53	1.23	54.25	**100**	37.00	2.35	60.65	**100**
Others	44.18	1.31	54.53	**100**	36.97	2.3	60.73	**100**
Type of house hold(HH)					**Type of house hold(HH)**			
Self-employed	47.44	1.09	51.48	**100**	38.83	1.86	59.31	**100**
Regular wage /salaried employee	45.49	1.58	52.94	**100**	40.79	2.52	56.69	**100**
Casual labour	46.38	1.23	52.41	**100**	41.44	2.04	56.52	**100**
Other	0.02	3.09	96.91	**100**	0.08	6.15	93.77	**100**
Head of HH					**Head of HH**			
Female self headed HH	52.74	0.21	47.06	**100**	33.35	0.75	65.9	**100**
Others	45.11	1.31	53.6	**100**	37.56	2.53	59.92	**100**
Number of Year of Formal Education					**Number of Year of Formal Education**			
0	37.23	0.05	62.74	**100**	19.79	0.21	80.01	**100**
1-4 year	25.26	0.08	74.67	**100**	16.61	0.22	83.17	**100**
5-8 year	49.26	0.49	50.26	**100**	37.91	0.74	61.36	**100**
9-10 year	57.62	1.26	41.13	**100**	46.34	1.38	52.29	**100**
11-12 year	52.59	2.56	44.87	**100**	40.8	2.72	56.49	**100**
13-17 year	55.44	8.48	36.09	**100**	49.39	7.69	42.93	**100**
>= 18 year	74.06	11.4	14.56	**100**	68.28	7.82	23.89	**100**
Marital status					**Marital status**			
Never married	16.12	2.37	81.53	**100**	18.03	4.47	77.51	**100**
Currently married	71.58	0.44	27.99	**100**	55	1	44.01	**100**
Widowed	40.04	0.07	59.91	**100**	26.86	0.23	72.91	**100**
Divorce/Separated	74.98	1.24	23.8	**100**	64.93	2.8	32.21	**100(**

Source: PLFS unit level data 2022-23

In both areas, we have seen that as the educational level increases beyond the 4 years level of formal education employment falls up to 11-12 years of formal education. Beyond that level of education employment increases and reaches a maximum at greater than 18 years of formal education (Table 1). A pertinent point to be noted is that quite a handful proportion of illiterate women in rural areas join as workers. More generally they are doing the low-paid job for necessary conditions. On the other hand, we see as educational level increases, beyond 11-12 years proportion of women seeking or available for job increases in both areas and it is more in rural areas. That may be compared to urban areas, rural educated females are more unable to find appropriate jobs.

Next, we consider the variable marital status of women to see the effect of this on women's activity status. In both areas, WPRs are greater among divorced/separated women followed by widows. However, these proportions are greater in rural areas compared to urban counterparts. The results reveal the fact that women who are divorced or widowed may have to bear more household responsibilities. Due to that reason, they are more likely to join as workers (Table 1). However unemployed proportion is greater among unmarried women compared to other marital statuses in both areas. The possible reason is that for unmarried women it is less costly to remain unemployed than for other marital status. They always try to find out better job.

5. Analytical Results:

5.1. **Variable used in multinomial Logistic Regression Analysis:**

Based on our descriptive analysis we have selected certain variables that may be important in explaining women's employment.

- **<u>Region:</u>**

A region is associated with different states that are more or less homogeneous. In the ***Northern region,*** we have considered Haryana, Himachal Pradesh (HP), Punjab, Delhi, Jammu & Kashmir(J&K); ***Western region*** consists of Maharashtra (MH), Gujarat, Goa, Lakshadweep, Damn & Diu.; ***Southern region*** consists of Kerala, Karnataka, Tamil Nadu(TN), Andhra Pradesh (AP), Puducherry, Andaman & Nicobor Islands (A&N); ***Eastern region*** consists of Assam, Odisha, West Bengal(WB), Northeastern region consist of Meghalaya, Mizoram, Manipur.

Based on poor demographic conditions, we have separated states like Bihar, Chhattisgarh, Madhya Pradesh (MP), Rajasthan, Uttar Pradesh (UP), and Uttarakhand as **EAG** states. India is heterogeneous in characteristics. A proper understanding of the disaggregate nature of women employment all India picture needs to be disaggregated at state level.

- **<u>Religion and social category:</u>**

In India, there are different religious people and also within each religion, there are different social categories. They are not homogeneous in terms of their culture, education, or socio-economic condition. Even among Muslims, there is also a social stratification in two broad categories such as Ashraf and Ajlaf based on Clans. It seems that a major portion of the Ajlaf community is Untouchable, backwards, engaged in low-paid occupations, and remains socioeconomically and educationally backwards. (Bharat Ch. Rout, 2017). To understand this heterogeneity in nature regarding women's employment we carried out our analysis among different religions and social groups.

- **<u>Household (HH) type:</u>**

People engaged broadly into three categories of employment Self-employment (SE), Regular Wage/salaried employees (RWE), and Casual labourers (CL). It seems that RWE is better than the other two categories in terms of salaried/wages, job security etc. Based on the main source of HH income, HH is categorized into SE, RWE, and CL types of HH. We want to see which type of HH affects women's engagement in the labour market.

- **<u>Household (HH) size:</u>**

A woman's employment also depends on the HH size. As HH size increases it may increase the burden of HH chores, and take care of responsibility that restricts them from moving out in the labour market.

- **<u>Marketed value of HH monthly per capita expenditure:</u>**

HH monthly per capita expenditure is used as a proxy of HH income. From lower family income women generally tend to enter into the labour market in a distressful job. As incomes go up the need for this distressed job reduces. Hence, they withdraw from the labour market.

- **<u>Head of the household:</u>**

If women become head of the HH then it will be more conducive for women to enter the job market. So we consider this variable to see the effect on the decision of women's labour market participation.

- **Education:**

Education has an important role in participating in the labour market. Illiterate and level literate people generally get offended jobs where whereas literate person tends to get decent jobs. So education is necessary to decide whether women enter in labour market or not.

We have considered education not as a continuous variable but at various levels according to the standard of formal education.

- **Age :**

We have considered the age of the working women who have some skills. Age is used as a proxy for their work experience. Here we consider only the working age group (15-59 years). People generally find better job opportunities up to a certain age because their work experience grows with the increase in age up to a certain level. Beyond that level, they cannot put full effort into work so they find it difficult to get the job.

We have considered three categories of female workers' age.: young (15-29), middle (30-40) and old age (40-59).

- **Marital status:**

Women's participation in the labour market is also determined by their marital status. A married woman has to bear a lot of HH chores, and childbearing activity than unmarried women. So it is expected married women are less likely to participate in the labour market than unmarried women. On the other hand, when women become widows or divorced, they have to bear the HH responsibility hence it is expected their participation in the labour market is greater.

5.2. Model of Multinomial logistic regression:

Now we discuss the basic structure of the regression model. If we have n independent observations with p-explanatory variables, and the categorical response variable has k categories, to construct the logits in the multinomial case, one of the categories must be considered the base level and all the logits are constructed relative to it. Any category can be taken as the base level, so we will take category k as the base level. Let π_j denote the multinomial Probability of an observation falling in the j^{th} category, to find the relationship between this probability and the p explanatory variables, $X_1, X_2, \ldots, X_p$, the multiple logistic regression model then is

$$\mathbf{Log}\ [\pi_j (x_i)/ \pi_k (x_i)] = \alpha_i + \beta_{1j}\, x_{1i} + \beta_{2j}\, x_{2i} + \ldots \beta_{pj} x_{pi}$$

Where j= 1, 2,…,(k-1), i= 1,2,….,n. Since all the π's add to unity, this reduces to

$$\mathbf{Log}\ (\pi_j\ (x_i)) = \exp\ (\alpha_i + \beta_{1j}\ x_{1i} + \beta_{2j}\ x_{2i} + \ldots \beta_{pj} x_{pi}) / \ 1 + \sum_{j=1}^{k-1} (\exp\,(\alpha_i + \beta_{1j}\, x_{1i} + \beta_{2j}\ x_{2i} + \cdots + \beta_{pj} x_{pi})$$

For j= 1, 2,…,(k-1), the model parameters are estimated by the method of ML. Practically, we use statistical software to do this fitting(Chatterjee and Hadi,2006; Habil,2012)

5.3: Results of Multinomial Approach:

A multinomial logistic regression model is utilized to examine the factors influencing women's engagement in the job market. This approach involves three categorical dependent variables such as employed, unemployed and not in the labour force, where not in the labour force is a base category. Table 2 contains the marginal effect derived from the multinomial logit model. The results suggest that the factors influencing the probability of women's participation in the labour market. For multinomial regression, marginal effects are more important than simple coefficients (Sengupta and Seth 2021).

The variables are chosen to keep in view the effects of various economic and non-economic factors which are responsible for women's participation in the labour market. Women's participation in the labour market is not only determined by economic factors but also by various socio-economic conditions. Various studies show women face various social and patriarchal constraints to move to the labour market. They are bound to do HH chores and take care responsibility of elders and children which is why they get less opportunity to develop their skills which leads to less opportunity to join the labour market. Along with these, it is societies' thinking that men will become the breadwinners and females will manage HH chores, for this reason, girls are neglected from their childhood to get opportunities in education, and skill enhancement programmes. But effects of these factors are not the same across all regions, religions and social groups across the whole country. Our country is heterogeneous in characteristics. In some societies, some cast, and some religions, women are freer to make the decision and can move out from their HH boundary. Women's labour supply is also affected by the income effect of the family. There is debate on how far education affects women's employment. Some argue that women and their families use higher education as marriage prospects. Others opine that quality of education enhances their earning capacity and opportunity to get better quality of job. In this study, we wish to test how far this logic carries to present PLFS data.

For variable HH size, we have seen that in rural areas household size negatively influences the probability of women's participation in the labour market as workers. In these areas, large families generally join the family. Women have lots of familial responsibility. Also, the family has many sources of earning. This may preclude women from joining the labour market. Along with these social customs, beliefs also play a role in this scenario.

The influence of the base category (not in the labour force) can often be derived indirectly. **If household size negatively affects "Worker" status, it suggests that individuals in larger households may be more likely to fall into other categories, such as "Not in the labour force" or "Seeking a job."**

In the urban area, surprisingly we also find that the probability of a fall in women's participation as workers with an increase in household size. In Indian urban areas also there is existence of join family. They are often engaged in various types of family businesses and or small-scale enterprises. The success of this depends on the bond and ties among the family members and their joint efforts. Thus the women may have a lesser need to go in the job market. Also, traditional values and customs often dominate traditional business and small-scale familial enterprises, making it very difficult for women to be engaged in the job market.

In urban areas, the negative effect on worker status and the positive effect on job-seeking suggest that people from larger households are less likely to be working and more likely to be looking for a job. Many of them may simply stay out of the labour force altogether.

Compared to other households we see that among the female self-headed household in the rural areas, the probability of women's participation in the labour market is higher. Generally, these household heads are either widows or separated from their husbands (either legally or through social custom). Hence they generally do not have any males within the working sector to find livelihood for them. So they are bound to offer their labour services to earn a living. In rural areas, there is a demand for labour in various agricultural and non-agricultural activities.

Urban areas also depict a higher probability of women working when they become the family head. The picture is the same as rural but there is magnitudinal difference between rural and urban areas. In urban areas headed households may include some spinsters. Also, they generally have more amount of human capital than rural sisters. These women may have some other ways of earning either through the possession of capital and or some assets. Consequently, the time they are also taking the training. Sometimes their participation in the labour market is not similar to that of rural areas. In rural areas, their participation is led by dire necessity due to a lack of other sources of income. In urban areas, it may be a more self-made decision depending on various opportunities and ways out. However, the pressure of maintaining the family still remains among the female-headed household. So we get the same positive effect on the probability of women's participation in the labour market from female-headed households in both areas.

The results show that women in female self-headed households are more likely to be working but less likely to be looking for a job. In rural areas, women in these households are more likely to be workers, meaning they are already employed or self-employed. However, the negative effect on job-seeking suggests that they are not actively searching for a new job because they are already working. In short, women in self-headed households are active in the workforce, but they may not be looking for a new job because they are already engaged in some form of work.

It has been also observed that for both the young and middle age groups their probability of participation in the labour market (either as employed or unemployed) increases at a certain level of age, beyond that age it reduces suggested by the variable square of age. In rural areas, the intuition behind the results may be that the majority of women are engaged in activities like agriculture and non-agricultural jobs, where the demand for physical labour is greater. Along with this their nutritional dietary level is also lower compared to their urban counterpart. So they can put full physical effort up to a certain age. Beyond that level of age, they are unable to put and hence it is difficult to find a job.

In urban areas, we also get similar results to rural areas in the relationship between age and the probability of women's work participation in the job market. However, the intuition behind the result is somewhat different from rural areas. In urban areas generally, women are engaged in various service sectors like teachers, nurses, paid household workers, clerks and so on. These jobs are smoothly conducted by their work experience. Hence with increasing age, their work experience is grown up. With the growth of experience, women find better opportunities for jobs up to a certain age. Beyond that they are unable to put in full effort and hence find it difficult to get a job.

So the relationship between age and women's participation in the labour market shows an inverted U hypothesis in both areas.

The positive effects for all age groups in both worker and job-seeking categories suggest that the base category mostly includes dependents like children and retirees. In rural areas, the young group shows high rates of job-seeking and worker, showing greater labour force activity than older individuals, implying younger household members are less represented in the base category.

For variable education, we have seen that in rural areas as the educational level among women increases then the probability of being a worker increases. A similar result is also demonstrated in urban areas.

People with low or no education are much more likely to be in the base category (not in the labour force). For example, in rural areas, having no formal education strongly reduces the chances of being a worker. As education increases, the negative impact on the worker and job-seeking categories diminishes, reflecting a lower likelihood of being entirely out of the labour force.

Monthly per capita expenditure calculated in Market value is used as a proxy of the income of the family. In rural areas where we see that as HH income increases there is less likelihood of women joining the labour market as workers. This might be because of generally rural women are generally forced to join the job market from less affluent families. They generally engaged in low-paid low-skilled jobs. As families move to higher and higher per capita income basket, the need for jobs sharply reduces. Hence, they are pulling out of the informal labour market.

In the urban area, we get similar results to rural areas for variable monthly per capita expenditure in the effect of the probability of women being workers. However, the marginal effect of a fall in women being workers with an increase in expenditure is less compared to rural areas. This leads to the fact that in urban areas women from affluent families also join in the labour market to a greater extent than in rural areas. Women from these families have better opportunities to gain more human development capital like better education, skills, and health status. So they get better opportunities to be engaged in the labour market. However social taboos and customs often limit their movement in the job market.

In urban areas, negative effects on worker status for the richer quintiles show that wealthier people are more likely to be in the base category (not in the labour force). In

rural areas, positive effects for the poorest quintiles on worker status mean that people from poorer households are more likely to be working.

Compared to other types of HH we have observed that SE, RWE, and CL types of HH have positive effects on women workers in rural areas. Women generally tend to join the labour market from this type of HH. Again comparing these three types of HH category we have seen that rural women are more likely to join in the labour market as workers from SE and CL types of HH(probability of 89%). This reflects the possible fact that when a household runs a small enterprise or business it needs more helping hand. They engaged female members of the family as helpers. In rural areas more generally women are engaged as helpers though they do not earn any independent income. Women in rural areas are also forced to engage in economic activity from casual types of HH the reason that generally casual types of households have less earning capacity.

In the urban area, we also get a similar effect to rural areas in the affection of types of HH on their employability condition. However, the magnitude of the probability of being a worker from these types of HH is greater in urban areas than that of in rural counterparts. The reason may be that urban women are generally better educated, and skilled than their rural counterparts. The availability of jobs for these women is also greater. Along with this patriarchy constraints affect less for their movement in the labour market.

Positive effects for worker status among SE (self-employed), RWE (regular wage earners), and CL (casual laborers) households suggest that households are less likely to have members in the base category (not in the labour force). Conversely, rural SE workers have a strong positive effect on worker status. At the same time, the negative effect on job-seeking shows that these households are also less likely to have members actively job-seeking, suggesting more stable employment.

In India, women belong to different religious groups such as Hindu, Muslim, and Christian, and so on. Similarly, women also belong to different social groups such as SC, ST, OBC, and other forward classes. In each religious group and social group is separated from others by different socio-economic and cultural backgrounds. We hope this will affect women's employability condition. Considering another caste as reference categories women from the Hindu Upper caste have a negative effect of being a worker while Muslim Unreserved are more likely to join in labour market as workers. This may lead to the fact that women of unreserved category belong to wealthy families. Women from these families have the better opportunity to gain a better level of education and skill. So they have better options to engage in the labor market as workers.

On the other hand in urban areas, women from Muslim Unreserved communities are more likely to engage as workers. This might imply the reason is that unreserved families have better options to gain better human capital resources. So they can easily enjoy their economic independence.

The effects for Hindu upper-caste individuals are negative, though not significant for job-seeking. This suggests that they are more likely to be in the base category (not in the labor force) compared to other religious groups. For individuals from the unreserved Muslim category, there is a negative effect on both worker and job-seeking status in rural areas, indicating they may be overrepresented in the base category. This suggests that they are less likely to be working or actively seeking jobs compared to other groups. Christian individuals show mixed effects. In rural areas, they have a positive effect on job-seeking, indicating they are more likely to be actively looking for work. In urban areas, the effect is negative on job-seeking, suggesting that they may be underrepresented among active job seekers in these areas. In simple terms, the results suggest that religion has an impact on whether individuals are in the labour force, with some religious groups being more likely to be in the base category (not in labour force) and others more actively engaged in the labour market.

Similarly, compared to other forward classes, women are more likely to join as workers from lower cast categories (ST, SC, and OBC) in rural areas. Socially backward classes are generally poor. To lead a smooth life it is necessary for them to join the labour market. It is also noticed that social taboo is not a constraint for the lower caste women. They generally enter in low level of work with a lower payment. However not working increases the social status of women in higher caste. In urban areas only from the ST community women are more likely to enter into the labour market as workers.

In rural areas, Scheduled Tribe (ST) individuals are more likely to work, meaning they are less represented in the base category (not in the labor force). However, they are less likely to look for jobs, suggesting their engagement in work reduces the likelihood of being in the job-seeking category. Scheduled Caste (SC) individuals are slightly less likely to be workers and job seekers, indicating they are more represented in the base category compared to other social groups. For Other Backward Classes (OBC), the differences are minor, meaning they are only slightly more likely to fall into the base category compared to other groups. In summary, individuals from Scheduled Tribes (ST) are more likely to be workers but less likely to be actively seeking jobs. Scheduled Castes (SC) and Other Backward Classes (OBC) show lower participation in the labor force and job-seeking, with OBC individuals showing the least difference from other groups

Next, we consider the variable region of India. A region is composed of different more or less homogeneous states. Across different regions, their socio-economic condition is also different. We wish to see this reflection on women's employability conditions across regions. Compared to the Northeastern region, women are more tend to be employed from EAG, Northern, Western, and Southern in rural areas. People in the EAG region are socially and economically distressed. They have fewer human development parameters in different dimensions. So it may be that females are forced to enter into the work to lead their lives properly. Again compared to different regions, we have seen that the probability of women being workers is greater in the Western region. This might be because these regions are more economically and socially developed. Social spending in this region is greater. This may influence the women's persuasion of better education and health. Also, Southern states are richer in respect of their culture. It gives the importance of women's employment.

In urban areas, women are more likely to participate as workers from Western and Eastern states. Again compared to these states we also have seen that the probability of women being workers is greater in Western but compared to rural areas it is much lower. It is also high in the Eastern region though it was negative in rural areas.

Regions like the Western area show a significant positive effect on worker status and negative effects on job-seeking, meaning fewer people are in the base category in these active economic regions. In the EAG states, there are mixed results—rural areas have more workers, but urban areas have fewer job-seekers, reflecting regional disparities in labour force participation.

Next, we wish to see the effect of the marital status of women on their participation in labour market as workers. In both rural and urban areas the relation is negative for unmarried but positive for married and divorced/separated. Widow and divorced/separated women more generally have to bear the economic responsibility of the family if there is no earning male member of the family. So they are compelled to engage in economic activity. Similarly it is true for married women. On the other hand, unmarried women face more security concerns issues to move into the labour market.

Unmarried women are less likely to work, meaning they are more likely to be in the base category, possibly as dependents. Married women are more likely to work or be job seekers, so they are less likely to be in the base category.

This section presents a detailed analysis of factors influencing women's labour market participation, using a multinomial logistic regression model with the dependent variable categorized as "employed," "unemployed," and "not in labour force" (base category). Key findings, derived from marginal effects, highlight how economic and non-economic factors shape women's labour supply, shaped by social norms, familial roles, and regional diversity.

Table 2: Marginal effects from the multinomial logit model for women (15-59 years) status

Multinomial logistic regression:

Relative probability of activity status of women

For rural area,
Number of obs= 70,047
LR chi2 (48) = 132089.46
Prob > chi2= 0.0000
Pseudo R2= 0.3614
Log likelihood= -116717.96

For urban area
Number of obs= 52,717
LR chi2 (48) =73805.26
Prob > chi2= 0.0000
Pseudo R2= 0.2750
Log likelihood= -97285.08

Dependent variable= worker, seeking or available for the job; Base category=not in labour force

Explanatory variables	Rural		Urban	
	Worker	Seeking or available for job	Worker	Seeking or available for job
1. HH size	-.0883264***	-.0016209	-.1084899***	.0724487***
	(.0028627)	(.0105668)	(.003302)	(.0083262)
2.Age categories				
a). Young	1.436695 ***	5.011639***	1.42147***	4.329101***
	(.0181732)	(.2995057)	(.0218897)	(.1927866)
b). Middle age	2.162452***	4.672872 ***	2.268968***	4.298038 ***
	(.0194659)	(.3044123)	(.0226768)	(.1955997)
c). Old age	2.336974***	3.279468***	2.229705***	3.122603 ***
	(.0181114)	(.3341831)	(.0217121)	(.2060375)

3.Percapita expenditure quintiles :	other(Ref.)			
a). Quintile1(poorer)	1.021765 (.4775301)	.3233061 (.7989757)	-.3072794 (.2194325)	13.54111 (616.7976)
b). Quintile2(middle)	.794199 (.4776765)	.3317655 (.7993317)	-.4605375* (.2196696)	13.31227 (616.7976)
c). Quintile3(richer)	.799146 (.4840336)	.1714067 (.8264884)	-.577438* (.2250602)	12.92245 (616.7976)
b). Quintile4(richest)	.280604 (.5438277)	-.315019 (1.090609)	-.2686209 (.2550938)	12.88087 (616.7978)
4. Female self-headed HH (other as Ref.)	.5700711*** (.0390516)	-1.533678*** (.2647232)	.3133097*** (.0420358)	-1.23083*** (.146246)
5. Number of Year of Formal Education (other as Ref.)				
a). Absence of formal education	-1.673065*** (.1190241)	-4.614805*** (.2413258)	-1.930776*** (.0606567)	-3.317436*** (.1590282)
b). Primary education	-1.731049*** (.1199242)	-4.122628*** (.26279)	-1.782376*** (.0635325)	-3.048474 *** (.1960542)
c). Upper Primary	-1.394581*** (.1186133)	-3.580203*** (.162537)	-1.370504*** (.0591432)	-2.844251*** (.1116057)
d).Maddhayamik	-1.245952*** (.1185637)	-3.436774*** (.1570531)	-1.298993*** (.0588198)	-2.882926*** (.1037162)
e). H.S	-1.229265*** (.1189464)	-2.99643*** (.1562041)	-1.370595*** (.0593462)	-2.53461*** (.1005963)
f). Graduation	-.8987659*** (.1192877)	-1.542682 (.1527014)	-.8655338*** (.0584475)	-1.072809*** (.0932105)
6. HH types: other(Ref.)				
a) SE	9.678637*** (1.000488)	-.851367*** (.0848716)	7.15758*** (.3543628)	-.8352555*** (.0576397)
b) RWE	9.396579*** (1.000531)	-.8728862*** (.0898698)	7.090436*** (.3543101)	-.634831*** (.054059)
c) CL	9.614941*** (1.000531)	-.5021189*** (.0917187)	7.387549*** (.3547667)	-.34012*** (.0773068)
7. Caste: other (Ref.)				
a) Hindu upper caste	-.0949053 (.04864)	-.2474891 (.1402184)	-.0705484 (.0440649)	-.2258538 (.1174631)
b) Islam unreserved	-.1403588** (.0518473)	-.6550378*** (.1614971)	.0140931 (.0480211)	-.1226194 (.1288714)
c) Christianity	-.0475987 (.0273864)	.3259175*** (.0844857)	.0167528 (.0321831)	-.0804228 (.0810984)
8. Social group: other(Ref.)				
a) ST	.5492514*** (.0485102)	-.0462011 (.1410673)	.1432292** (.0474478)	.0897913 (.1256339)

b) SC	-.0420056	-.1595609	.1585118***	.1466952
	(.0483364)	(.1389963)	(.0452759)	(.1201144)
c) OBC	-.0065872	-.1521759	.0460775***	-.0981085
	(.0475197)	(.135845)	(.0436088)	(.1159836)
9. Region: North eastern(Ref.)				
a) EAG	.1550512 ***	-.4831049***	-.0900026***	-.2647888***
	(.022845)	(.0771765)	(.0250348)	(.0672187)
b) Northern	.3501787***	.2146856 *	-.0425689	-.0430679
	(.0264991)	(.084885)	(.0278501)	(.0744813)
c) Western	.4041205 ***	-.3750403***	.1054356**	-.4293702***
	(.0269614)	(.0965071)	(.0272249)	(.0784332)
d) Southern	.2027849***	-.1505665	-.009328	-.1659576*
	(.0250446)	(.0813492)	(.0255271)	(.0689587)
e) Eastern	-.0334251	-.1603853	.0603603*	-.1371477
	(.0265691)	(.0900611)	(.0305508)	(.0856922)
10. Marital status: widowed (Ref.)				
a) Unmarried	-1.248546***	.3616075	-.4970526***	.5785601**
	(.029842)	(.349406)	(.0370933)	(.2210619)
b) Married	.5462389***	. -.0787585	.3221926***	-.3014116
	(.028723)	(.3479106)	(.0354705)	(.218656)
c) Divorce/separated	.8104854***	1.132218**	.8842092***	.974598***
	(.100389)	(.4765872)	(.0867727)	(.2986944)

Note: Standard errors in parentheses; *** significance at 1% level, ** significance at 5% level, *significance at 10% level; Ref: reference category; HH: household; SC: Scheduled Caste; ST: Scheduled Tribe; OBC: other backward caste

Source: Source: authors' calculation from PLFS 2022-23

6. Conclusion:

The study attempts to find out the factors that influence women's activity status such as employed, unemployed, and not in the labour force. First, it attempts to find out some salient features of such women's activity status across different variables. The analysis shows that women's activity status more specifically women's employment varies across region, religion, social group, and marital status. The entire analysis is depends on rural and urban differently. In urban areas, irrespective of each variable female employment is always lower than that of rural regions. This reveals the possible fact that by nature women in urban areas prefer to remain domestic domain, do HH duty, childbearing, and rearing activities, where the income effect is much stronger than in rural areas. On the other hand, compared to urban areas, HH income in rural areas is generally lower, so women have to join as workers to support their families.

However irrespective of each variable proportion of women's unemployment is greater in urban areas than rural areas but educated women's unemployment is greater in rural areas than in urban areas. This clearly reveals that rural-educated females cannot find appropriate jobs more than the urban area.

The result found that more women from developed regions such as Northern, Southern, and Western regions and underdeveloped regions like EAG in rural areas come to join the Labour market Women whereas from Eastern region women are less likely to participate as workers. In rural areas, women from the Hindu upper caste have a relatively positive effect on women's employment whereas in urban areas women from the Christian community have a positive effect on women's employment. Women from lower social groups such as ST are more likely to join as workers. The relationship between age and women's employment has an inverted U-shaped relationship whereas the relationship between education and employment has U shaped relationship. We also find out that women from more affluent families are less likely to participate in the labour market. Also result shows that women who have decision-making power, more of them come to join as workers. Unmarried and divorced women are more likely to join in labour market than married women.

Most of the variables affect women's employment in the same dimension in both rural and urban areas but their marginal effect varies in both areas. Where we see women in urban areas, to get a job, education, and skill is more important than experience. In urban areas, social thinking about women is much more developed than in rural areas. So women face less constraint by patriarchal society in urban areas. They have more advantages to move in the labour market than in rural areas.

One of the important aspects of female workers is that education in no way enhances the possibility of female employment. This is a very sorry state of affairs. This clearly indicates a cleavage between educational skills and employability skills. It is said that in India, formal education does not necessarily create the possibility of being employed. It is high time that government should take proper steps in skill enhancement. It is of utmost importance to improve the employability of women. Also, social awareness regarding women's employment is to be propagated by the government.

7. Policy implications:

Based on the findings of the study, several policy implications can be suggested to enhance women's participation in the labor force and reduce inequalities between rural and urban areas, as well as across social and educational groups. Policy implications are as follows-

1. Skill Development Programs: The study highlights that education alone does not guarantee women’s employment opportunities, indicating a skills gap where practical skills are lacking despite having formal education. The government should prioritize vocational training and skill development programs for women, especially in rural areas, focusing on market-relevant skills in high-employment sectors like technology, healthcare, retail, and manufacturing. Specialized re-skilling programs should be implemented for educated but unemployed women to align their qualifications with the job market's needs.

2. Increasing Awareness and Changing Social Norms:- The study found that women in urban areas have greater access to the labor market, likely due to less patriarchal constraints compared to rural areas. Cultural and social norms still limit their participation in certain sectors in urban

areas. The government should invest in awareness campaigns to challenge outdated gender norms and promote the value of women's work in both rural and urban areas. These initiatives should also stress the advantages of women’s workforce participation, such as economic independence and social progress. In rural areas, where traditional gender roles are more prominent, policies should promote community-based programs that involve both women and men in conversations about shared responsibilities in household work and women’s economic contributions.

3. Encouraging Female Employment in Rural Areas:- The study shows that women in rural areas are more likely to join the labor market due to lower household income, but they also face more barriers to employment than their urban counterparts. The government should focus on creating local job opportunities in rural areas, particularly for women. This could be achieved by promoting small-scale industries, self-employment initiatives, and expanding agricultural and rural development programs that generate jobs for women. Furthermore, policies aimed at improving rural infrastructure, such as enhanced access to transportation, markets, and financial services, could help more women enter and remain in the workforce.

4. Support for Unmarried, Divorced, and Widowed Women:- The study indicates that unmarried and divorced women are more inclined to join the labor force than married women, possibly due to a greater need for financial independence or lack of caregiving responsibilities. The government should implement social safety nets for women, particularly those from disadvantaged backgrounds (e.g., unmarried, divorced, widowed), to provide the necessary support for balancing work and family duties. Policies like flexible working hours, childcare assistance, and legal safeguards against discrimination will help motivate women to stay in the workforce, particularly in urban areas.

5. Incentivizing Women’s Participation in the Formal Sector:- Many women are employed in the informal sector, which offers limited job security, benefits, and fair wages. Policies that encourage women to move into the formal sector could enhance their working conditions and financial stability. The government could provide incentives to companies that hire women, especially in sectors like manufacturing, technology, and finance. Furthermore, tax reductions or subsidies for businesses that offer flexible work arrangements, equal pay, and family-friendly policies could also promote increased female participation in formal employment.

6. Improving Access to Decision-Making and Leadership Roles: The study reveals that women who have decision-making authority within households or communities are more likely to join the workforce. Empowering women to make decisions both at home and in the workplace is essential for achieving long-term gender equality in employment. Policies should promote female leadership and involvement in decision-making at all levels—whether in households, businesses, or governments. This could involve backing women’s entrepreneurship initiatives, offering leadership training, and providing mentorship opportunities to help women move into management or decision-making positions.

7. Targeted Interventions for Specific Social Groups:-The study finds that women from certain social groups (e.g., ST women, women from Hindu upper castes, or Christian communities) have different rates of employment, highlighting the need for more inclusive policies. There is a need for focused initiatives that tackle the unique challenges faced by marginalized groups of women, such as those from Scheduled Tribes (ST) or minority communities. These initiatives could aim at overcoming barriers to education, skill development, and employment, while also addressing discrimination and social exclusion.

8. Improving the Quality of Education and Link to Employment:- As noted, education alone does not guarantee better employment opportunities for women. This disparity indicates a gap between educational attainment and actual employability. The government should revise educational curricula to focus on vocational training, technical education, and internship opportunities. Collaborations with private industries can help ensure that educational institutions provide women with the skills necessary to meet labor market requirements. Moreover, improving access to secondary and higher education, particularly for rural women, is key to enhancing employment opportunities, but it should also include career counseling and assistance with the transition from education to employment.

To bridge the gaps in women's labor force participation, policies should emphasize skills development, raising social awareness, improving infrastructure, and creating job opportunities, especially for rural and marginalized women. By crafting inclusive and context-driven strategies, the government can foster a more fair labor market that provides women with enhanced opportunities for economic independence and social advancement.

Acknowledgment

The authors would like to thank anonymous referees and the Editorial Advisory Board (EAB) of this journal for their valuable suggestions at different stages of revision of this paper.

References:

1. Andres, Luis Alberto; Dasgupta, Basab; Joseph, George; Abraham, Vinoj and Correia, Maria C.(2017): "Precarious drop: Reassessing Patterns of Female Labor Force Participation in India", World Bank Policy Research, Working Paper 8024
2. Azzez Abdul E P,Dandub Palzor Negi, Asha Rani & Senthil Kumar A P .(2021): "The Impact of Covid-19 on Migrant Women Workers in India",Eurasian Geography And Econmics,2021,Vol.62,No.1,93-112.
3. Basu Kaushik, Sanghamitra Das and Bhaskar Dutta (2010): " Child Labour and Household Wealth: Theory and Empirical Evidence of an Inverted –U", Journal of Developmentnomics ,91,2010, 8-14.
4. Bergman BR (1974): "Occupational segregation, wages and profits when employers discriminate by race and sex", Eastern Economic Journal 1:103–110 April–July.
5. Brinton, M.C., Lee, Y., and Parish, W.L. (1995): "*Married women's employment in rapidly industrializing societies: Examples from East Asia*", American Journal of Sociology 100(5): 1099–1130. doi:10.1086/230633
6. Chatterjee,S and A.Hadi (2006):Regression Analysis by Example,John Wiley & Sons.
7. Chattopadhya Raghabendra and Esther Duflo (2004): "Women as Policy Makers: Evidence from a randomized policy Experiment in India", Econometrica, Econometric Society, Vol.72(5), pages 1409-1443,September.
8. Chaudhary, Ruchika and Sher Verick (2014): *"Female Labour Force Participation in India and Beyond"* , ILO Asia-Pacific working paper series,ISSN 2227-4391;2227-4405
9. Dhanaraj, Sowmya; Mahambare, Vidya (2017) : "*Family structure, education and women's employment in rural India"*, WIDER Working Paper, No. 2017/195, ISBN 978-

92-9256-421-6, The United Nations University World Institute for Development Economics Research (UNU-WIDER), Helsinki, https://doi.org/10.35188/UNU-WIDER/2017/421-6

10. Drèze, J., and A. Sen. (2013): *"An Uncertain Glory: India and its contradictions*", London: A. Lane.
11. Engels Friedrich (2010) : The Origin Of The Family ,Private Property And The State,Penguin classic, Reissue Edition ,ISBN-10 : 0141191112
12. Ghai, Surbhi (2018): "The Anomaly of Women's Work and Education in India", Indian Council For Research On International Economic Relations, Working Paper 368, December 2018
13. Habib ,Irfan(2007) : "*Medieval India: The Study of a Civilization",*
14. Habil,Abdalla M. El(2012): " An Application on Multinomial Logistic Regression Model",Pakistan Journal of Statistics and Operation Research,March 2012,
15. Hirway, I (2012): "*Missing Labour Force: An Explanation*, "Economic & Political Weekly, Vol 47,No 37, pp 67–72
16. Jajoria, Deepika and Manoj Jadav(2020): " *Is Periodic Labour Force Survey,2017-18 Comparable with Employment –Unemployment Survey,2011-12?",* Economic & Political Weekly,VOL LV NO 3, 2020
17. Kabeer, N. (2012). Women's economic empowerment and inclusive growth: Labour markets and enterprise development. International Development Research Centre (IDRC) and Department for International Development (DFID).
18. Mehrotra ,Santosh and Sharmistha Sinha(2017): *" Explaining Falling Female Employment during a High Growth Period",* Economic & Political Weekly,VOL LII NO 39, 2017
19. Mehrotra Santosh and Jajati K Parida.(2021): "*Stalled Structural Change Bringsan Employment Crisis in India"* The Indian Journal of Labour Economics,2021,64:281-308. https://doi.org/10.1007//s41027-021-00317-2-x.
20. Mill John Stuart (1869): The Subjection of Women,Longmans,Green, Reader and Dyer, London, 1869
21. Mies Maria (1982): Lace Makers of Narsapur: Indian Housewives Produce for the World Market (Spinifex Feminist Classics)
22. Mitra,Arup;Guru Prakash Singh and Puneet Kumar shrivastav (2020): *" How Unstable the Sources of Livelihood Are? Analysis on Periodic Labour Force Survey",* working paper number : SAUFE-WP- 2020-008
23. Nanda, P., Datta, N., & Das, P. (2014). *Son preference and daughter neglect in India: What happens to living girls?* International Center for Research on Women (ICRW).
24. Papola ,T S(2014): " An Assessment of The Labour StatisticsSystem in India," ILO country office for India,New Delhi
25. PROBE (1999): *"Public Report on Basic Education in India*", Oxford University Press, New Delhi
26. Reed N. Megan (2020): "The labour Forcer Participation of Tinian women before and after widowhood", Demographic Research, vol.43, Article 24,pp 673-706.
27. Rustagi Preet (2005): "Understanding Gender Inequalities in Wages and Income in India", the Indian Journal of Labour Economics, Vol.48, No.2, 2005.

28. Sarkar, Sudipa,Soham Sahoo and Stephan Klasen(2018): "Employment Transition of Women in India: A panel Analysis" The University of Warwick , https://wrap.warwick.ac.uk/112260
29. Sen Amartya (1984): Resource ,Values and Development, OUP, New Delhi.
30. Sen Amartya (1987): 'The Standard of Living: Lecture II. Lives and Capabilities in Sen, Muelebauer ,Kanbur, Hart and Williams ,The Standard of Living, The Tanner Lectures, Cambridge University Press, Cambridge.
31. Sen Amartya (1992): Inequality Reexamined, Russell Sage Foundation, Clarendon Press,Oxford , New York.
32. Sen Amartya(2001): " Many Faces Gender Inequality", based on inauguration lecture for the New Radcliff Institute at Harvard University ,2001
33. Sengupta, Atanu and Tushar Mukherjee(2010): Firms and Their Female Labourers-An Indian Informal Industry: a Case Study of 'BIDI' making in Murshidabad District in West Bengal , India ,LAP LAMBERT Academic Publishing,2010
34. Sengupta ,Atanu and Ujjwal Seth (2021): " *A Subjective Approach of Firm's Performance: A Multinomial Analysis*", SARVEKSHANA,ISSN 2249-197X, 110 & 111, pp 58-69.
35. Shrivastav(2024) *"Understanding Migration Behaviour in the Indian from PLFS Data Post-lockdown and Future Perspective",* Economics and Political Weekly , Vol 57, No 53, pp22-24.
36. Sundari S. (2020): *"Structural Changes and Quality of Women's Labour in India*", The Indian Journal of Labour Economics, 2020, 63:689-717. https://doi.org/10.1007//s41027-020-00245-2
37. Swaminathan, Madhura(2020): *"Contemporary Features of Rural Workers in India with Focus on Gender and Caste",*The Indian Journal of Labour Economics. https://doi.org/10.1007//s41027-020-00210-z.
38. Tilak, Jandhyala B.G. (2002): *''Education Poverty in India''*, Review of Development and Change, Vol. 7, No. 1.
39. Wollstonecraft Mary (1792): A Vindication of the Rights of Women: With Strictures of Political and Moral Subjects, Boston, Faust Statue No. 45, Newbury- Street, 1792.

PART-II

Highlights of Reports Released by National Statistics Office (NSO)

(The 'Highlights' are reproduced from related report prepared by Household Survey Division (HSD) of NSO. For details, the reader may refer to the related main report.)

SARVEKSHANA

Highlights of Recent Survey Report(s) Released by National Statistics Office (NSO)

1. Periodic Labour Force Survey (PLFS) (2023-24)

HIGHLIGHTS

Periodic Labour Force Survey (PLFS) 2023-2024

Survey Period

July 2023 to June 2024

Survey Coverage

Surveyed

12,743 First Stage Units (FSUs)	Rural: 6,975 villages Urban: 5,768 urban blocks
1,01,920 Households	55,796 in rural areas 46,124 in urban areas
4,18,159 Persons	2,42,546 in rural areas 1,75,613 in urban areas

The survey covered the whole of the Indian Union *except* the villages in Andaman and Nicobar Islands which remained extremely difficult to access throughout the year.

Approaches for presenting Labour Force Indicators

Approaches followed for presenting Labour Force Indicators

usual status (ps+ss) Reference period : 1 year	current weekly status(CWS) Reference period : 1 week

Some of the key results at the all-India level for the period July 2023 - June 2024 emerging from PLFS are highlighted below.

A. Labour Force in usual status (ps+ss)

LFPR
for persons
of all ages

Labour Force Participation Rate (LFPR) in usual status (ps+ss) in India: 45.1%

rural	urban
male: 57.9%	male: 59.0%
female: 35.5%	female: 22.3%

LFPR
for persons
of age 15-29 years

Labour Force Participation Rate (LFPR) in usual status (ps+ss) for persons of age 15-29 years in India: 46.5%

rural:	urban:
48.1%	42.6%

LFPR
for persons
of age 15 years and above

Labour Force Participation Rate (LFPR) in usual status (ps+ss) for persons of age 15 years and above in India: 60.1%

rural:	urban:
63.7%	52.0%

B. Workforce

WPR
in usual
status for
persons of
all ages

Worker Population Ratio (WPR) in usual status (ps+ss) in India: 43.7%

rural	urban
male: 56.3% female: 34.8%	male: 56.4% female: 20.7%

WPR
in usual
status for
persons of
age 15-29
years

Worker Population Ratio (WPR) in usual status (ps+ss) for persons of age 15-29 years in India: 41.7%

rural:	urban:
44.0%	36.3%

WPR
in usual
status for
persons of
age 15
years and
above

Worker Population Ratio (WPR) in usual status (ps+ss) for persons of age 15 years and above in India: 58.2%

rural:	urban:
62.1%	49.4%

Status in employment among workers in usual status (ps+ss)

Share (%) of self-employed among workers in usual status (ps+ss)

rural male: 59.4	rural female:73.5	urban male: 39.8	urban female:42.3

Share (%) of regular wage/ salaried employees among workers in usual status (ps+ss)

rural male:15.8	rural female: 7.8	urban male: 46.8	urban female: 49.4

Share (%) of casual labour among workers in usual status (ps+ss)

rural male:24.9	rural female: 18.7	urban male:13.4	urban female:8.3

C. Unemployment Rate in usual status (ps+ss)

Unemployment Rate (UR) in usual status for persons of all ages

Unemployment Rate in usual status (ps+ss) for persons of all age in India: 3.2%

rural male: 2.7% female: 2.1%	urban male: 4.4% female: 7.1%

Unemployment Rate (UR) in usual status for persons of age 15 years and above

Unemployment Rate in usual status (ps+ss) for persons of age 15 years and above in India: 3.2%

rural male: 2.7% female: 2.1%	urban male: 4.4% female: 7.1%

Unemployment Rate (UR) in usual status for educated persons of all age 15 years and above

Unemployment Rate in usual status (ps+ss) for educated (*highest level of education secondary and above*) persons of age 15 years and above in India:7.1%

rural	urban
6.5%	7.9%

Unemployment Rate (UR) in usual status for persons of age 15 -29 years

Unemployment Rate in usual status (ps+ss) for youth persons of age 15 -29 years in India: 10.2%

rural	urban
male: 8.7%	male:12.8%
female: 8.2%	female: 20.1%

D.Time Series of Key Labour Force indicators in usual status (ps+ss) obtained from PLFS

Table 1: Labour force participation rates (in per cent) in usual status (ps+ss) estimated from PLFS (2017-18), PLFS(2018-19), PLFS (2019-20), PLFS (2020-21), PLFS (2021-22), PLFS (2022-23) and PLFS (2023-24) for persons of age 15 years and above and persons of all ages

all-India

age group	rural			urban			rural+urban		
	male	female	person	male	female	person	male	female	person
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PLFS (2023-24)									
15 years and above	80.2	47.6	63.7	75.6	28.0	52.0	78.8	41.7	60.1
all ages	57.9	35.5	46.8	59.0	22.3	41.0	58.2	31.7	45.1
PLFS (2022-23)									
15 years and above	80.2	41.5	60.8	74.5	25.4	50.4	78.5	37.0	57.9
all ages	55.5	30.5	43.4	58.3	20.2	39.8	56.2	27.8	42.4
PLFS (2021-22)									
15 years and above	78.2	36.6	57.5	74.7	23.8	49.7	77.2	32.8	55.2
all ages	56.9	27.2	42.2	58.3	18.8	39.0	57.3	24.8	41.3
PLFS (2020-21)									
15 years and above	78.1	36.5	57.4	74.6	23.2	49.1	77.0	32.5	54.9
all ages	57.1	27.7	42.7	58.4	18.6	38.9	57.5	25.1	41.6
PLFS (2019-20)									
15 years and above	77.9	33.0	55.5	74.6	23.3	49.3	76.8	30.0	53.5
all ages	56.3	24.7	40.8	57.8	18.5	38.6	56.8	22.8	40.1
PLFS (2018-19)									
15 years and above	76.4	26.4	51.5	73.7	20.4	47.5	75.5	24.5	50.2
all ages	55.1	19.7	37.7	56.7	16.1	36.9	55.6	18.6	37.5
PLFS (2017-18)									
15 years and above	76.4	24.6	50.7	74.5	20.4	47.6	75.8	23.3	49.8
all ages	54.9	18.2	37.0	57.0	15.9	36.8	55.5	17.5	36.9

2023-24 refers to the period July 2023 – June 2024 and likewise for 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 and 2017-18

Table2: WPR (in per cent) in usual status (ps+ss) estimated from PLFS (2017-18), PLFS(2018-19), PLFS (2019-20), PLFS (2020-21), PLFS (2021-22), PLFS (2022-23) and PLFS (2023-24) for persons of age 15 years and above and persons of all ages

all-India

age group	rural			urban			rural+urban		
	male	female	person	male	female	person	male	female	person
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PLFS (2023-24)									
15 years and above	78.1	46.5	62.1	72.3	26.0	49.4	76.3	40.3	58.2
all ages	56.3	34.8	45.6	56.4	20.7	38.9	56.4	30.7	43.7
PLFS (2022-23)									
15 years and above	78.0	40.7	59.4	71.0	23.5	47.7	76.0	35.9	56.0
all ages	54.0	30.0	42.3	55.6	18.7	37.7	54.4	27.0	41.1
PLFS (2021-22)									
15 years and above	75.3	35.8	55.6	70.4	21.9	46.6	73.8	31.7	52.9
all ages	54.7	26.6	40.8	55.0	17.3	36.6	54.8	24.0	39.6
PLFS (2020-21)									
15 years and above	75.1	35.8	55.5	70.0	21.2	45.8	73.5	31.4	52.6
all ages	54.9	27.1	41.3	54.9	17.0	36.3	54.9	24.2	39.8
PLFS (2019-20)									
15 years and above	74.4	32.2	53.3	69.9	21.3	45.8	73.0	28.7	50.9
all ages	53.8	24.0	39.2	54.1	16.8	35.9	53.9	21.8	38.2
PLFS (2018-19)									
15 years and above	72.2	25.5	48.9	68.6	18.4	43.9	71.0	23.3	47.3
all ages	52.1	19.0	35.8	52.7	14.5	34.1	52.3	17.6	35.3
PLFS (2017-18)									
15 years and above	72.0	23.7	48.1	69.3	18.2	43.9	71.2	22.0	46.8
all ages	51.7	17.5	35.0	53.0	14.2	33.9	52.1	16.5	34.7

2023-24 refers to the period July 2023 – June 2024 and likewise for 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 and 2017-18

Table3: Unemployment Rate (in per cent) in usual status (ps+ss) estimated from PLFS (2017-18), PLFS(2018-19), PLFS (2019-20), PLFS (2020-21), PLFS (2021-22), PLFS (2022-23) and PLFS (2023-24) for persons of age 15 years and above and persons of all ages

all-India

age group	rural			urban			rural+urban		
	male	female	person	male	female	person	male	female	person
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PLFS (2023-24)									
15 years and above	2.7	2.1	2.5	4.4	7.1	5.1	3.2	3.2	3.2
all ages	2.7	2.1	2.5	4.4	7.1	5.1	3.2	3.1	3.2
PLFS (2022-23)									
15 years and above	2.7	1.8	2.4	4.7	7.5	5.4	3.3	2.9	3.2
all ages	2.8	1.8	2.4	4.7	7.5	5.4	3.3	2.9	3.2
PLFS (2021-22)									
15 years and above	3.8	2.1	3.2	5.8	7.9	6.3	4.4	3.3	4.1
all ages	3.8	2.1	3.3	5.8	7.9	6.3	4.4	3.3	4.1
PLFS (2020-21)									
15 years and above	3.8	2.1	3.3	6.1	8.6	6.7	4.5	3.5	4.2
all ages	3.9	2.1	3.3	6.1	8.6	6.7	4.5	3.5	4.2
PLFS (2019-20)									
15 years and above	4.5	2.6	3.9	6.4	8.9	6.9	5.0	4.2	4.8
all ages	4.5	2.6	4.0	6.4	8.9	7.0	5.1	4.2	4.8
PLFS (2018-19)									
15 years and above	5.5	3.5	5.0	7.0	9.8	7.6	6.0	5.1	5.8
all ages	5.6	3.5	5.0	7.1	9.9	7.7	6.0	5.2	5.8
PLFS (2017-18)									
15 years and above	5.7	3.8	5.3	6.9	10.8	7.7	6.1	5.6	6.0
all ages	5.8	3.8	5.3	7.1	10.8	7.8	6.2	5.7	6.1

2023-24 refers to the period July 2023 – June 2024 and likewise for 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 and 2017-18

E. Time Series of Key Labour Force indicators in Current Weekly Status (CWS) obtained from PLFS

Table 1: Labour force participation rates (in per cent) current weekly status estimated from PLFS (2017-18), PLFS(2018-19), PLFS (2019-20), PLFS (2020-21), PLFS (2021-22), PLFS (2022-23) and PLFS (2023-24) for persons of age 15 years and above and persons of all ages

all-India

age group	rural			urban			rural+urban		
	male	female	person	male	female	person	male	female	person
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PLFS (2023-24)									
15 years and above	78.7	39.7	58.9	75.0	26.1	50.8	77.5	35.6	56.4
all ages	56.7	29.6	43.2	58.5	20.8	40.0	57.3	27.1	42.3
PLFS (2022-23)									
15 years and above	78.8	34.6	56.7	73.9	24.0	49.4	77.4	31.6	54.6
all ages	54.5	25.4	40.4	57.9	19.1	39.0	55.4	23.7	40.0
PLFS (2021-22)									
15 years and above	76.7	29.2	53.0	74.2	22.1	48.6	75.9	27.2	51.7
all ages	55.7	21.7	38.9	57.9	17.5	38.2	56.3	20.5	38.7
PLFS (2020-21)									
15 years and above	76.7	30.0	53.4	73.8	21.7	48.0	75.8	27.5	51.8
all ages	56.0	22.7	39.7	57.8	17.3	38.0	56.5	21.2	39.2
PLFS (2019-20)									
15 years and above	76.7	28.3	52.5	73.8	22.1	48.2	75.8	26.3	51.2
all ages	55.4	21.1	38.6	57.2	17.5	37.8	56.0	20.0	38.3
PLFS (2018-19)									
15 years and above	75.5	22.5	49.1	73.7	19.7	47.1	74.9	21.6	48.5
all ages	54.5	16.7	36.0	56.7	15.6	36.7	55.2	16.4	36.2
PLFS (2017-18)									
15 years and above	75.6	21.7	48.9	74.1	19.6	47.1	75.1	21.1	48.4
all ages	54.4	16.1	35.7	56.7	15.3	36.4	55.0	15.8	35.9

2023-24 refers to the period July 2023 – June 2024 and likewise for 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 and 2017-18

Table 2: Worker Population Ratio (in per cent) current weekly status estimated from PLFS (2017-18), PLFS(2018-19), PLFS (2019-20), PLFS (2020-21), PLFS (2021-22), PLFS (2022-23) and PLFS (2023-24) for persons of age 15 years and above and persons of all ages

all-India

age group	rural			urban			rural+urban		
	male	female	person	male	female	person	male	female	person
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PLFS (2023-24)									
15 years and above	75.3	38.1	56.5	70.5	23.9	47.4	73.8	33.8	53.7
all ages	54.3	28.4	41.4	55.0	19.0	37.3	54.5	25.7	40.2
PLFS (2022-23)									
15 years and above	75.2	33.2	54.2	69.3	21.8	46.0	73.5	30.0	51.8
all ages	52.0	24.4	38.6	54.2	17.4	36.3	52.6	22.5	38.0
PLFS (2021-22)									
15 years and above	71.7	27.9	49.9	68.4	19.9	44.6	70.7	25.6	48.3
all ages	52.1	20.7	36.6	53.4	15.7	35.0	52.4	19.3	36.1
PLFS (2020-21)									
15 years and above	71.2	28.6	50.0	66.8	19.0	43.1	69.9	25.7	47.9
all ages	52.0	21.6	37.1	52.4	15.2	34.1	52.1	19.8	36.3
PLFS (2019-20)									
15 years and above	70.1	26.7	48.4	66.0	19.4	43.0	68.8	24.4	46.7
all ages	50.6	19.9	35.5	51.2	15.4	33.6	50.8	18.6	35.0
PLFS (2018-19)									
15 years and above	69.0	20.9	45.0	67.2	17.4	42.7	68.4	19.8	44.3
all ages	49.7	15.5	32.9	51.7	13.7	33.2	50.3	15.0	33.0
PLFS (2017-18)									
15 years and above	69.1	20.1	44.8	67.7	17.1	42.6	68.6	19.2	44.1
all ages	49.6	14.8	32.6	51.7	13.3	32.9	50.2	14.4	32.7

2023-24 refers to the period July 2023 – June 2024 and likewise for 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 and 2017-18

Table 3: Unemployment Rate (in per cent) current weekly status estimated from PLFS (2017-18), PLFS(2018-19), PLFS (2019-20), PLFS (2020-21), PLFS (2021-22), PLFS (2022-23) and PLFS (2023-24) for persons of age 15 years and above and persons of all ages

all-India

age group	rural			urban			rural+urban		
	male	female	person	male	female	person	male	female	person
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PLFS (2023-24)									
15 years and above	4.4	3.9	4.2	6.0	8.7	6.7	4.8	5.0	4.9
all ages	4.4	3.9	4.2	6.0	8.7	6.7	4.9	5.0	4.9
PLFS (2022-23)									
15 years and above	4.6	4.0	4.4	6.3	9.1	7.0	5.1	5.1	5.1
all ages	4.6	4.0	4.5	6.3	9.1	7.0	5.1	5.1	5.1
PLFS (2021-22)									
15 years and above	6.5	4.5	6.0	7.8	9.9	8.3	6.9	5.8	6.6
all ages	6.5	4.6	6.0	7.8	9.9	8.3	6.9	5.8	6.6
PLFS (2020-21)									
15 years and above	7.1	4.8	6.5	9.4	12.2	10.1	7.8	6.6	7.5
all ages	7.2	4.8	6.5	9.4	12.2	10.1	7.8	6.6	7.5
PLFS (2019-20)									
15 years and above	8.7	5.5	7.8	10.5	12.4	11.0	9.3	7.3	8.8
all ages	8.7	5.5	7.9	10.6	12.4	11.0	9.3	7.3	8.8
PLFS (2018-19)									
15 years and above	8.6	7.3	8.3	8.8	12.1	9.5	8.7	8.7	8.7
all ages	8.7	7.3	8.4	8.9	12.1	9.5	8.8	8.7	8.8
PLFS (2017-18)									
15 years and above	8.7	7.5	8.4	8.7	12.7	9.5	8.7	9.0	8.7
all ages	8.8	7.7	8.5	8.8	12.8	9.6	8.8	9.1	8.9

2023-24 refers to the period July 2023 – June 2024 and likewise for 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 and 2017-18

खण्ड-III हिंदी

सर्वेक्षण

राष्ट्रीय सांख्यिकी कार्यालय की पत्रिका

भाग सं० PDOS-57-XXXX सितंबर 2024 (1-2)
अंक संख्या 117वां
सितंबर, 2024

सत्यमेव जयते

राष्ट्रीय सांख्यिकी कार्यालय
सांख्यिकी और कार्यक्रम कार्यान्वयन मंत्रालय
भारत सरकार
नई दिल्ली

एनएसओ द्वारा जारी की गई रिपोर्ट की मुख्य बातें

(मुख्य बातें एनएसओ के एच.एस.डी. प्रभाग द्वारा तैयार की गई सम्बंधित रिपोर्ट से उद्धृत की गई हैं। विवरण के लिए पाठक सम्बंधित मुख्य रिपोर्ट देख सकते हैं)

सर्वेक्षण
भाग सं० PDOS-57-XXXX सितंबर 2024 (1-2)

मुख्य बातें

आवधिक श्रमबल सर्वेक्षण (पीएलएफएस) 2023-24

सर्वेक्षण अवधि

जुलाई 2023 से जून 2024

सर्वेक्षण कवरेज

सर्वेक्षण किया गया

12,743 फर्स्ट स्टेज उनिट्स (एफएसयु)	ग्रामीण: 6,975 गांवों नगरीय: 5,768 नगरीय खंडो
1,01,920 परिवारों	55,796 ग्रामीण क्षेत्रों में 46,124 नगरीय क्षेत्रों में
4,18,159 वयक्तियों	2,42,546 ग्रामीण क्षेत्रों में 1,75,613 नगरीय क्षेत्रों में

इस सर्वेक्षण में पूरे भारतीय संघ को शामिल किया गया अंडमान और निकोबार द्वीप समूह के उन गाँवों को छोड़कर जिन तक पहुँच पाना पूरे वर्ष तक बेहद कठिन था।

श्रम बल संकेतक पेश करने के आधार

श्रम बल संकेतक पेश करने के लिए अपनाया गया आधार

सामान्य स्थिति (पीएस+एसएस) सन्दर्भ अबधि: 1 वर्ष	वर्तमान साप्ताहिक स्थिति(सीडब्ल्यूएस) सन्दर्भ अबधि: 1 सप्ताह

अखिल भारतीय स्तर पर जुलाई 2023 - जून 2024 की अवधि के लिए पीएलएफएस से प्राप्त कुछ मुख्य परिणाम निम्नलिखित हैं। (क)

(क) श्रमबल सामान्य स्थिति (पीएस+एसएस) में

एलएफपीआर
सभी उम्र के
व्यक्तियों में

भारत में लेबर फोर्स पार्टिसपैशन रेट (एलएफपीआर) सामान्य स्थिति (पीएस+एसएस) में: 45.1%

ग्रामीण	नगरीय
पुरूषों में: 57.9%	पुरूषों में: 59.0%
महिलाओं में: 35.5%	महिलाओं में: 22.3%

एलएफपीआर
15-29 वर्ष उम्र
के व्यक्तियों
में

भारत के 15-29 वर्ष उम्र के व्यक्तियों में सामान्य स्थिति (पीएस+एसएस) में एलएफपीआर: 46.5%

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
48.1%	42.6%

(ख) कार्यबल

डब्ल्यूपीआर
सभी उम्र के
व्यक्तियों में

कामगार जनसंख्या अनुपात (डब्ल्यूपीआर) सामान्य स्थिति में (पीएस+एसएस) में: 43.7%

ग्रामीण	नगरीय
पुरूषों में: 56.3% महिलाओं में: 34.8%	पुरूषों में: 56.4% महिलाओं में: 20.7%

डब्ल्यूपीआर
15-29 वर्ष उम्र
के व्यक्तियों
में

15-29 वर्ष उम्र के व्यक्तियों में भारत में डब्ल्यूपीआर सामान्य स्थिति (पीएस+एसएस) में: 41.7%

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
44.0%	36.3%

डब्ल्यूपीआर
15 वर्ष एवं
उससे अधिक
उम्र के
व्यक्तियों में

15 वर्ष एवं उससे अधिक उम्र के व्यक्तियों में भारत में डब्ल्यूपीआर सामान्य स्थिति (पीएस+एसएस) में: 58.2%

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
62.1%	49.4%

सामान्य स्थिति (पीएस+एसएस) में कामगारों के बीच रोजगार स्थिति

सामान्य स्थिति (पीएस+एसएस) में कामगारों के बीच स्व-रोजगार का शेयर(%)

ग्रामीण	ग्रामीण	नगरीय	नगरीय
पुरूषों में: 59.4	महिलाओं में: 73.5	पुरूषों में: 39.8	महिलाओं में: 42.3

सामान्य स्थिति (पीएस+एसएस) में कामगारों के बीच नियमित मजदूर/वेतनभोगी कर्मचारियों का शेयर(%)

ग्रामीण	ग्रामीण	नगरीय	नगरीय
पुरूषों में: 15.8	महिलाओं में: 7.8	पुरूषों में: 46.8	महिलाओं में: 49.4

सामान्य स्थिति (पीएस+एसएस) में कामगारों के बीच आकस्मिक मजदूरों का शेयर(%)

ग्रामीण	ग्रामीण	नगरीय	नगरीय
पुरूषों में: 24.9	महिलाओं में: 18.7	पुरूषों में: 13.4	महिलाओं में: 8.3

(ग) बेरोजगार दर सामान्य स्थिति (पीएस+एसएस) में

बेरोजगार दर सभी उम्र के व्यक्तियों पर

भारत में सामान्य स्थिति (पीएस+एसएस) में बरोजगार दर: 3.2%

ग्रामीण	नगरीय
पुरूषों में: 2.7% महिलाओं में: 2.1%	पुरूषों में: 4.4% महिलाओं में: 7.1%

बरोजगार दर 15 वर्ष एवं उससे अधिक उम्र के

भारत में सामान्य स्थिति (पीएस+एसएस) में बरोजगार दर 15 वर्ष एवं उससे अधिक उम्र के व्यक्तियों में: 3.2%

ग्रामीण	नगरीय
पुरूषों में: 2.7% महिलाओं में: 2.1%	पुरूषों में: 4.4% महिलाओं में: 7.1%

बरोजगार दर 15 बर्षो और उससे उपर के उम्र के शिक्षित

भारत में 15 बर्षो और उससे उपर के उम्र के शिक्षित (माध्यमिक एवं उसके उच्चतर का अधिकतम शिक्षा का स्तर) व्यक्तियों में सामान्य स्थिति (पीएस+एसएस) में बरोजगार दर: 7.1%

ग्रामीण क्षेत्रों में	नगरीय क्षेत्रों में
6.5%	7.9%

बरोजगार दर युवा (15-29 वर्ष उम्र के) व्यक्तियों में

भारत में बरोजगार दर युवा (15-29 वर्ष उम्र के) व्यक्तियों में: 10.2%

<u>ग्रामीण</u>	<u>नगरीय</u>
पुरूषों में: 8.7%	पुरूषों में: 12.8%
महिलाओं में: 8.2%	महिलाओं में: 20.1%

(घ). प्रमुख श्रम बल संकेतक की टाइम-सीरीज सामान्य स्थिति (पीएस+एसएस) में पीएलएफएस से प्राक्कलित

टेबल 1: लेबर फोर्स पार्टिसपैशन रेट (एलएफपीआर) (प्रतिशत में) सामान्य स्थिति (पीएस+एसएस) में पीएलएफएस (2017-18), पीएलएफएस (2018-19), पीएलएफएस (2019-20), पीएलएफएस (2020-21), पीएलएफएस (2021-22), पीएलएफएस (2022-23) एवं पीएलएफएस (2023-24) से प्राक्कलित

आल-इंडिया

आयु वर्ग	ग्रामीण			नगरीय			ग्रामीण + नगरीय		
	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
पीएलएफएस (2023-24)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	80.2	47.6	63.7	75.6	28.0	52.0	78.8	41.7	60.1
सभी उम्र के व्यक्तियों के लिए	57.9	35.5	46.8	59.0	22.3	41.0	58.2	31.7	45.1
पीएलएफएस (2022-23)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	80.2	41.5	60.8	74.5	25.4	50.4	78.5	37.0	57.9
सभी उम्र के व्यक्तियों के लिए	55.5	30.5	43.4	58.3	20.2	39.8	56.2	27.8	42.4
पीएलएफएस (2021-22)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	78.2	36.6	57.5	74.7	23.8	49.7	77.2	32.8	55.2
सभी उम्र के व्यक्तियों के लिए	56.9	27.2	42.2	58.3	18.8	39.0	57.3	24.8	41.3
पीएलएफएस (2020-21)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	78.1	36.5	57.4	74.6	23.2	49.1	77.0	32.5	54.9
सभी उम्र के व्यक्तियों के लिए	57.1	27.7	42.7	58.4	18.6	38.9	57.5	25.1	41.6
पीएलएफएस (2019-20)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	77.9	33.0	55.5	74.6	23.3	49.3	76.8	30.0	53.5
सभी उम्र के व्यक्तियों के लिए	56.3	24.7	40.8	57.8	18.5	38.6	56.8	22.8	40.1
पीएलएफएस (2018-19)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	76.4	26.4	51.5	73.7	20.4	47.5	75.5	24.5	50.2
सभी उम्र के व्यक्तियों के लिए	55.1	19.7	37.7	56.7	16.1	36.9	55.6	18.6	37.5
पीएलएफएस (2017-18)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	76.4	24.6	50.7	74.5	20.4	47.6	75.8	23.3	49.8
सभी उम्र के व्यक्तियों के लिए	54.9	18.2	37.0	57.0	15.9	36.8	55.5	17.5	36.9

2023-24 जुलाई 2023-जून 2024 की अवधि को संदर्भित करता है और इसी तरह 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 और 2017-18 के लिए

टेबल 2: कामगार जनसंख्या अनुपात (डब्ल्यूपीआर) (प्रतिशत में) सामान्य स्थिति (पीएस+एसएस) में पीएलएफएस (2017-18), पीएलएफएस (2018-19), पीएलएफएस (2019-20), पीएलएफएस (2020-21), पीएलएफएस (2021-22), पीएलएफएस (2022-23) एवं पीएलएफएस (2023-24) से प्राक्कलित

आल-इंडिया

आयु वर्ग	ग्रामीण			नगरीय			ग्रामीण + नगरीय		
	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
पीएलएफएस (2023-24)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	78.1	46.5	62.1	72.3	26.0	49.4	76.3	40.3	58.2
सभी उम्र के व्यक्तियों के लिए	56.3	34.8	45.6	56.4	20.7	38.9	56.4	30.7	43.7
पीएलएफएस (2022-23)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	78.0	40.7	59.4	71.0	23.5	47.7	76.0	35.9	56.0
सभी उम्र के व्यक्तियों के लिए	54.0	30.0	42.3	55.6	18.7	37.7	54.4	27.0	41.1
पीएलएफएस (2021-22)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	75.3	35.8	55.6	70.4	21.9	46.6	73.8	31.7	52.9
सभी उम्र के व्यक्तियों के लिए	54.7	26.6	40.8	55.0	17.3	36.6	54.8	24.0	39.6
पीएलएफएस (2020-21)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	75.1	35.8	55.5	70.0	21.2	45.8	73.5	31.4	52.6
सभी उम्र के व्यक्तियों के लिए	54.9	27.1	41.3	54.9	17.0	36.3	54.9	24.2	39.8
पीएलएफएस (2019-20)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	74.4	32.2	53.3	69.9	21.3	45.8	73.0	28.7	50.9
सभी उम्र के व्यक्तियों के लिए	53.8	24.0	39.2	54.1	16.8	35.9	53.9	21.8	38.2
पीएलएफएस (2018-19)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	72.2	25.5	48.9	68.6	18.4	43.9	71.0	23.3	47.3
सभी उम्र के व्यक्तियों के लिए	52.1	19.0	35.8	52.7	14.5	34.1	52.3	17.6	35.3
पीएलएफएस (2017-18)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	72.0	23.7	48.1	69.3	18.2	43.9	71.2	22.0	46.8
सभी उम्र के व्यक्तियों के लिए	51.7	17.5	35.0	53.0	14.2	33.9	52.1	16.5	34.7

2023-24 जुलाई 2023-जून 2024 की अवधि को संदर्भित करता है और इसी तरह 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 और 2017-18 के लिए

टेबल 3: बरोजगार दर (प्रतिशत में) सामान्य स्थिति (पीएस+एसएस) में पीएलएफएस (2017-18), पीएलएफएस (2018-19), पीएलएफएस (2019-20), पीएलएफएस (2020-21), पीएलएफएस (2021-22), पीएलएफएस (2022-23) एवं पीएलएफएस (2023-24) से प्राक्कलित

आल-इंडिया

आयु वर्ग	ग्रामीण			नगरीय			ग्रामीण + नगरीय		
	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
पीएलएफएस (2023-24)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	2.7	2.1	2.5	4.4	7.1	5.1	3.2	3.2	3.2
सभी उम्र के व्यक्तियों के लिए	2.7	2.1	2.5	4.4	7.1	5.1	3.2	3.1	3.2
पीएलएफएस (2022-23)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	2.7	1.8	2.4	4.7	7.5	5.4	3.3	2.9	3.2
सभी उम्र के व्यक्तियों के लिए	2.8	1.8	2.4	4.7	7.5	5.4	3.3	2.9	3.2
पीएलएफएस (2021-22)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	3.8	2.1	3.2	5.8	7.9	6.3	4.4	3.3	4.1
सभी उम्र के व्यक्तियों के लिए	3.8	2.1	3.3	5.8	7.9	6.3	4.4	3.3	4.1
पीएलएफएस (2020-21)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	3.8	2.1	3.3	6.1	8.6	6.7	4.5	3.5	4.2
सभी उम्र के व्यक्तियों के लिए	3.9	2.1	3.3	6.1	8.6	6.7	4.5	3.5	4.2
पीएलएफएस (2019-20)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	4.5	2.6	3.9	6.4	8.9	6.9	5.0	4.2	4.8
सभी उम्र के व्यक्तियों के लिए	4.5	2.6	4.0	6.4	8.9	7.0	5.1	4.2	4.8
पीएलएफएस (2018-19)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	5.5	3.5	5.0	7.0	9.8	7.6	6.0	5.1	5.8
सभी उम्र के व्यक्तियों के लिए	5.6	3.5	5.0	7.1	9.9	7.7	6.0	5.2	5.8
पीएलएफएस (2017-18)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	5.7	3.8	5.3	6.9	10.8	7.7	6.1	5.6	6.0
सभी उम्र के व्यक्तियों के लिए	5.8	3.8	5.3	7.1	10.8	7.8	6.2	5.7	6.1

2023-24 जुलाई 2023-जून 2024 की अवधि को संदर्भित करता है और इसी तरह 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 और 2017-18 के लिए

(ङ). प्रमुख श्रम बल संकेतक की टाइम-सीरीज साप्ताहिक स्थिति (सीडब्ल्यूएस) में पीएलएफएस से प्राक्कलित

टेबल 1: लेबर फोर्स पार्टिसपैशन रेट (एलएफपीआर) (प्रतिशत में) वर्तमान साप्ताहिक स्थिति (सीडब्ल्यूएस) में पीएलएफएस (2017-18), पीएलएफएस (2018-19), पीएलएफएस (2019-20), पीएलएफएस (2020-21), पीएलएफएस (2021-22), पीएलएफएस (2022-23) एवं पीएलएफएस (2023-24) से प्राक्कलित

आल-इंडिया

आयु वर्ग	ग्रामीण			नगरीय			ग्रामीण + नगरीय		
	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
पीएलएफएस (2023-24)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	78.7	39.7	58.9	75.0	26.1	50.8	77.5	35.6	56.4
सभी उम्र के व्यक्तियों के लिए	56.7	29.6	43.2	58.5	20.8	40.0	57.3	27.1	42.3
पीएलएफएस (2022-23)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	78.8	34.6	56.7	73.9	24.0	49.4	77.4	31.6	54.6
सभी उम्र के व्यक्तियों के लिए	54.5	25.4	40.4	57.9	19.1	39.0	55.4	23.7	40.0
पीएलएफएस (2021-22)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	76.7	29.2	53.0	74.2	22.1	48.6	75.9	27.2	51.7
सभी उम्र के व्यक्तियों के लिए	55.7	21.7	38.9	57.9	17.5	38.2	56.3	20.5	38.7
पीएलएफएस (2020-21)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	76.7	30.0	53.4	73.8	21.7	48.0	75.8	27.5	51.8
सभी उम्र के व्यक्तियों के लिए	56.0	22.7	39.7	57.8	17.3	38.0	56.5	21.2	39.2
पीएलएफएस (2019-20)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	76.7	28.3	52.5	73.8	22.1	48.2	75.8	26.3	51.2
सभी उम्र के व्यक्तियों के लिए	55.4	21.1	38.6	57.2	17.5	37.8	56.0	20.0	38.3
पीएलएफएस (2018-19)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	75.5	22.5	49.1	73.7	19.7	47.1	74.9	21.6	48.5
सभी उम्र के व्यक्तियों के लिए	54.5	16.7	36.0	56.7	15.6	36.7	55.2	16.4	36.2
पीएलएफएस (2017-18)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	75.6	21.7	48.9	74.1	19.6	47.1	75.1	21.1	48.4
सभी उम्र के व्यक्तियों के लिए	54.4	16.1	35.7	56.7	15.3	36.4	55.0	15.8	35.9

2023-24 जुलाई 2023-जून 2024 की अवधि को संदर्भित करता है और इसी तरह 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 और 2017-18 के लिए

टेबल 2: कामगार जनसंख्या अनुपात (डब्ल्यूपीआर) (प्रतिशत में) साप्ताहिक स्थिति (सीडब्ल्यूएस) में पीएलएफएस (2017-18), पीएलएफएस (2018-19), पीएलएफएस (2019-20), पीएलएफएस (2020-21), पीएलएफएस (2021-22), पीएलएफएस (2022-23) एवं पीएलएफएस (2023-24) से प्राक्कलित आल-इंडिया									
आयु वर्ग	ग्रामीण			नगरीय			ग्रामीण + नगरीय		
	पुरुषों में	महिलाओं में	व्यक्तियों में	पुरुषों में	महिलाओं में	व्यक्तियों में	पुरुषों में	महिलाओं में	व्यक्तियों में
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
पीएलएफएस (2023-24)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	75.3	38.1	56.5	70.5	23.9	47.4	73.8	33.8	53.7
सभी उम्र के व्यक्तियों के लिए	54.3	28.4	41.4	55.0	19.0	37.3	54.5	25.7	40.2
पीएलएफएस (2022-23)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	75.2	33.2	54.2	69.3	21.8	46.0	73.5	30.0	51.8
सभी उम्र के व्यक्तियों के लिए	52.0	24.4	38.6	54.2	17.4	36.3	52.6	22.5	38.0
	पीएलएफएस (2021-22)								
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	71.7	27.9	49.9	68.4	19.9	44.6	70.7	25.6	48.3
सभी उम्र के व्यक्तियों के लिए	52.1	20.7	36.6	53.4	15.7	35.0	52.4	19.3	36.1
	पीएलएफएस (2020-21)								
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	71.2	28.6	50.0	66.8	19.0	43.1	69.9	25.7	47.9
सभी उम्र के व्यक्तियों के लिए	52.0	21.6	37.1	52.4	15.2	34.1	52.1	19.8	36.3
	पीएलएफएस (2019-20)								
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	70.1	26.7	48.4	66.0	19.4	43.0	68.8	24.4	46.7
सभी उम्र के व्यक्तियों के लिए	50.6	19.9	35.5	51.2	15.4	33.6	50.8	18.6	35.0
	पीएलएफएस (2018-19)								
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	69.0	20.9	45.0	67.2	17.4	42.7	68.4	19.8	44.3
सभी उम्र के व्यक्तियों के लिए	49.7	15.5	32.9	51.7	13.7	33.2	50.3	15.0	33.0
	पीएलएफएस (2017-18)								
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	69.1	20.1	44.8	67.7	17.1	42.6	68.6	19.2	44.1
सभी उम्र के व्यक्तियों के लिए	49.6	14.8	32.6	51.7	13.3	32.9	50.2	14.4	32.7
2023-24 जुलाई 2023-जून 2024 की अवधि को संदर्भित करता है और इसी तरह 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 और 2017-18 के लिए									

टेबल 3: बरोजगार दर (प्रतिशत में) साप्ताहिक स्थिति (सीडब्ल्यूएस) में पीएलएफएस (2017-18), पीएलएफएस (2018-19), पीएलएफएस (2019-20), पीएलएफएस (2020-21), पीएलएफएस (2021-22), पीएलएफएस (2022-23) एवं पीएलएफएस (2023-24) से प्राक्कलित

आल-इंडिया

आयु वर्ग	ग्रामीण			नगरीय			ग्रामीण + नगरीय		
	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में	पुरूषों में	महिलाओं में	व्यक्तियों में
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
पीएलएफएस (2023-24)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	4.4	3.9	4.2	6.0	8.7	6.7	4.8	5.0	4.9
सभी उम्र के व्यक्तियों के लिए	4.4	3.9	4.2	6.0	8.7	6.7	4.9	5.0	4.9
पीएलएफएस (2022-23)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	4.6	4.0	4.4	6.3	9.1	7.0	5.1	5.1	5.1
सभी उम्र के व्यक्तियों के लिए	4.6	4.0	4.5	6.3	9.1	7.0	5.1	5.1	5.1
पीएलएफएस (2021-22)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	6.5	4.5	6.0	7.8	9.9	8.3	6.9	5.8	6.6
सभी उम्र के व्यक्तियों के लिए	6.5	4.6	6.0	7.8	9.9	8.3	6.9	5.8	6.6
पीएलएफएस (2020-21)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	7.1	4.8	6.5	9.4	12.2	10.1	7.8	6.6	7.5
सभी उम्र के व्यक्तियों के लिए	7.2	4.8	6.5	9.4	12.2	10.1	7.8	6.6	7.5
पीएलएफएस (2019-20)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	8.7	5.5	7.8	10.5	12.4	11.0	9.3	7.3	8.8
सभी उम्र के व्यक्तियों के लिए	8.7	5.5	7.9	10.6	12.4	11.0	9.3	7.3	8.8
पीएलएफएस (2018-19)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	8.6	7.3	8.3	8.8	12.1	9.5	8.7	8.7	8.7
सभी उम्र के व्यक्तियों के लिए	8.7	7.3	8.4	8.9	12.1	9.5	8.8	8.7	8.8
पीएलएफएस (2017-18)									
15 वर्ष और उससे अधिक उम्र के व्यक्तियों के लिए	8.7	7.5	8.4	8.7	12.7	9.5	8.7	9.0	8.7
सभी उम्र के व्यक्तियों के लिए	8.8	7.7	8.5	8.8	12.8	9.6	8.8	9.1	8.9

2023-24 जुलाई 2023-जून 2024 की अवधि को संदर्भित करता है और इसी तरह 2022-23, 2021-22, 2020-21, 2019-20, 2018-19 और 2017-18 के लिए

BROAD CONTENTS OF THE LAST FOUR ISSUES OF 'SARVEKSHANA'

(A) 113th ISSUE

PART –I: TECHNICAL PAPERS

1. Estimating Annual Change in Inventory for the Unorganised Sector – an alternative approach

Dipanker Mitra

2. Status of Morbidity and Prevalence of Disease Specific Morbidity in Assam NSS 75th Round

Nayanakhee Sarma and Rajshree Bedamatta

3. Disaggregate Level Estimation and Mapping of Time Spent by Women on Unpaid Activities via Small Area Modeling by Integrating Survey and Census Data

Sonakhya Samaddar, Hukum Chandra

PART-II: HIGHLIGHTS OF THE REPORT RELEASED BY NSO:

4. Highlights I: Annual Report of Periodic Labour Force Survey (PLFS), July 2020-June 2021

5. Highlights II: Migration in India, 2020-2021

(B) 114th ISSUE

PART –I: TECHNICAL PAPERS

1. Making India Open Defecation Free: A Machine Learning Approach

Piyush Kumar Sing

2. Estimation of Quality Adjusted Life Year (QALY) based on Discrete Axiom of Revealed Preferences (DARP) and Splines for Different States of India

Gurprit Grover & Radhika Magan Bedamatta

3. Recent Estimates of Dynamic Mobility of Persons in Current Weekly Activity Status Based on Markov Chain

P. D. Joshi & Jena P Joshi

PART-II: HIGHLIGHTS OF THE REPORT RELEASED BY NSSO:

4. Highlights: Annual Report of Periodic Labour Force Survey (PLFS) (2021-2022)

(C) Combined (115th & 116th) ISSUE

PART –I: TECHNICAL PAPERS

1. Interstate Variations in Consumption Gap and Income Gap

P. D. Joshi & Jena P Joshi

2. Quantifying and Comparing Underemployment: A Multidimensional Approach using Fuzzy Logic

Sonakhya Samaddar

3. Decoding Assam's Crime Landscape: A Longitudinal Statistical Study

Puspita Saha & Dibyojyoti Bhattacharjee

PART-II: HIGHLIGHTS OF THE REPORT RELEASED BY NSSO:

4. Highlights: Annual Report of Periodic Labour Force Survey (PLFS) (2022-2023)

शुभकामनाओं सहित
महानिदेशक (रा. प्र.स.)
राष्ट्रीय सांख्यिकी कार्यालय,
सांख्यिकी एवं कार्यक्रम कार्यान्वयन मंत्रालय
भारत सरकार
सांख्यिकी भवन, महार्षि वाल्मीकि मार्ग,
सीबीडी शाहदरा, दिल्ली- 110032, भारत

With the compliments from
Director General (NSS)
National Statistics Office
Ministry of Statistics and Programme Implementation
Government of India
Sankhyiki Bhawan, Maharshi Valmiki Marg,
CBD Shahdara, Delhi-110032, INDIA

ACKNOWLEDGEMENT CARD

I acknowledge with thanks the receipt of publication 'Sarvekshana' Issue No. 117

I wish/do not wish to continue receiving the publication 'Sarvekshana'

(Signature) ....Name.............................
Address...................................
....

Date.................................

To,

Director General (NSS)
National Statistics Office
Coordination & Quality Control Division
Ministry of Statistics and Programme Implementation
Government of India
Sankhyiki Bhawan, Maharshi Valmiki Marg,
CBD Shahdara, Delhi-110032, INDIA

PDOS-57-XXXX Sept. 2024 (1-2)
305

The objective of this journal is to present analytical articles based on the results published on data collected through NSS, other surveys and censuses. 'Sarvekshana' welcomes contributions based on the NSS & other data sets and those relating to its methodology.

- Authors are requested to send soft copy of their papers/ articles in MS Word format to the Additional Director General, Coordination & Quality Control Division, National Statistics Office, Sankhyiki Bhawan, Maharshi Valmiki Marg, Near Karkardooma Court, Delhi–110032 through e-mail @ nssocpd.coord@mospi.gov.in.

- Papers/ articles should be of single-spaced with margins of 1½ " wide on the left and 1''on the right in MS Word format. The content of the article/paper should be in the sequence of **title of the paper, name of authors with their affiliations as foot note, abstract, keywords, main text part, acknowledgments and references.**

- For mathematical formulae, symbols, etc., standard notation may be followed. The graphs, charts and diagrams, if any, should be complete with all details and be such that direct reproduction can be obtained.

- The abstract should include summary of the paper not exceeding 150 words and a maximum of 6 important keywords used in the paper should also be mentioned.

- References should only include those papers, books etc. which have citation in the current paper and the same should be given in full e.g. name of author, title of publication, name of journal, date of publication etc.

- Subscription to 'Sarvekshana' should be sent to the Controller of Publications, Department of Publication, Civil Lines, Delhi-110054.

- Sarvekshana can also be downloaded free of cost from our website mospi.gov.in after due process of Internet enabled registration by the User/Researcher/Student.

Price: ₹300/-

SUGGESTIONS FOR IMPROVEMENT
ARE WELCOME