Vol. No. PDOS-XXXVIII No. (3-4)
ISSN 2249-197X

सर्वेक्षण

SARVEKSHANA

114th Issue
March, 2023

**Journal of
National Sample Survey Office**

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

SARVEKSHANA

114th Issue

Journal of
National Sample Survey Office

सत्यमेव जयते

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

Journal of National Sample Survey Office (NSSO), 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,
Survey Coordination Division
National Sample Survey Office, MoSPI
Sankhyiki Bhawan, Maharshi Valmiki Marg, Delhi – 110032

Foreword

Bringing out *Sarvekshana* has always been an enlightening endeavour. The first issue of *Sarvekshana* was released during July, 1977 and the latest released issue is 113th issue. The present 114th issue comes with three papers on the subjects of (i) Making India Open Defecation Free: A Machine Learning Approach, (ii) Estimation of Quality Adjusted Life Year (QALY) Based on Discrete Axiom of Revealed Preferences (DARP) and Splines for Different States of India and (iii) Recent Estimates of Dynamic Mobility of Persons in Current Weekly Activity Status Based on Markov Chain. In addition, the highlights of the recent survey report of Periodic Labour Force Survey namely, 'Annual Report of Periodic Labour Force Survey (PLFS), July 2021 – June 2022' have been included in the 114th issue.

Referees have been very 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 them and solicit continued support for the Journal. Authors of the papers have also been cooperative in accepting the suggestions for repetitive revisions. I congratulate them for their work which we hope would be useful. Officers of Survey Coordination Division of National Sample Survey 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.

G. C. Manna

New Delhi

April, 2023

Chairman

Editorial Advisory Board

Contents

PART-I

TECHNICAL PAPERS

Making India Open Defecation Free: A Machine Learning Approach

Piyush Kumar Sing[1]

Abstract

The objective of this paper is to create a classifier using machine learning algorithms that can identify individuals who do not use a latrine. If such individuals can be identified, then they may be encouraged to use latrines by appropriate policy interventions, thus making India open defecation free. For this purpose, 4 different models are created and the input variables for the first three models are chosen heuristically. But, the input variables for the fourth model are chosen from another machine learning model to remove any bias or prior knowledge of any variables. A comparison of all these models is done to choose the best model. Many such classifiers may exist but we are aiming for a very simple (having fewer input variables) and effective classifier.

Keywords: Machine Learning, Open Defecation Free, Latrine Usage, NSS data, India.

JEL Code: C38, C52, C53

Receipt of Final Version of paper from Author: April 2023

Acceptance: April 2023

[1] Piyush Kumar Sing, ISS, Deputy Director, DGCIS, Ministry of Commerce and Industry, Kolkata 700107. The views expressed, herein, are solely of the author and not of the organization/office where he is serving.

1. Introduction

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. It is also defined as an application of AI that enables systems to learn and improve from experience without being explicitly programmed. Machine learning focuses on developing computer programs that can access data and use it to learn for themselves. Through the use of statistical methods, algorithms are trained to make classifications or predictions, **uncovering key insights within data** mining projects. There are mainly four types of machine learning problems - supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. This paper focuses on a supervised learning problem using a machine learning algorithm known as CatBoost (Categorical Boosting) Classifier.

CatBoost provides a gradient boosting framework that among other features attempts to solve for categorical features using a permutation-driven alternative compared to the other gradient boosting algorithms. In prediction problems involving unstructured data (images, text, etc.) artificial neural networks tend to outperform all other algorithms or frameworks. However, when it comes to small-to-medium structured/tabular data, decision tree-based algorithms are considered best-in-class right now.

CatBoost is considered to be one of the best-performing machine learning algorithms for tabular data having a lot of categorical variables in the present day. NSS 76th round data contains a lot of categorical variables and that is one reason for choosing CatBoost for this paper. CatBoost can also rank the input variables according to their importance in explaining model behaviour, which is a very useful and unique feature. This is another reason for choosing this particular machine learning algorithm.

It has been widely established that poor sanitation and the practice of open defecation has disastrous impacts on the health of an individual and economies on a larger scale. A recent joint monitoring programme (JMP) on water, sanitation and hygiene by the World Health Organization and UNICEF released on July 1, 2021 stated that at least 15% of the total population in India defecates in the open. One per cent of the urban and 22% of rural population practises open defecation in the country (Shivangi, 2021).

Any country aims to eliminate the black spot of open defecation. The Government of India is trying its best to address this issue (through projects like Swachh Bharat Mission) but the problem always lies in identifying the ‘right’ person so that they can be motivated to use latrines. It is possible to make complex mathematical models which could identify these persons with a very high level of precision. Still, when these models are applied in the field, they are not found to be practical because they use too many inputs for prediction.

The aim of this paper is to identify individuals who do not use latrine using a ML model. To make such predictions we can make an infinite number of models using the same NSS data but in this paper, we are trying to make a simple (very less input variables) yet effective model.

There can be many factors which can contribute to latrine non-usage by an individual. For example, age can be a factor, level of education can be a factor etc. But the statement ‘age or level of education can contribute to latrine non usage’ is based on our experience and knowledge of life. This may or may not be true. It is possible that some variable ‘V1’ alone can make a better prediction than age and level of education combined.

In the first part of the paper, we built 3 models by using variables based on our knowledge and experience and as we increased the number of variables, we found that the model performance started to improve.

In the second part, we tried to make a model which does not have any knowledge bias. For this purpose, we choose variables which are contributing more to explain the variability as suggested by the machine (in the first part) are contributing more to explain the variability. When this is done, we landed with a model which not only has the least number of input variables but is also the most effective among the models we created in part one. This shows how ML algorithms when used properly can give out of the box solutions for complex problems.

2. Data

(A) Data Source

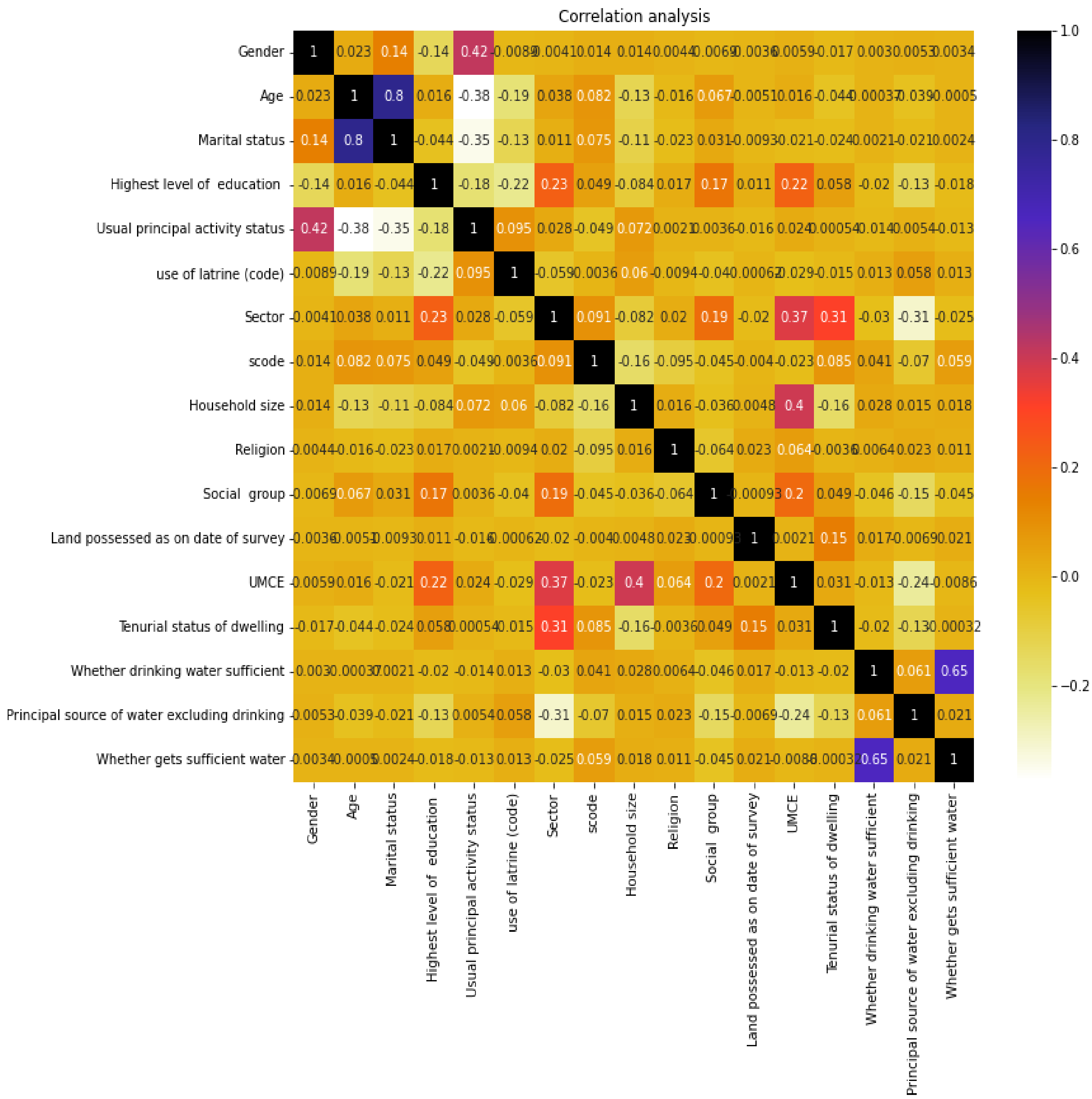
National Sample Survey Office (NSSO) conducted an all-India survey on 'Drinking Water, Sanitation, Hygiene and Housing Condition' in its NSS 76th round from July, 2018 to December, 2018. The objective of this survey was to examine and study different aspects of living conditions necessary for decent and healthy living. In this survey information was collected on access to latrine, in terms of exclusive use, common use or no access, type of latrine and reason for not using latrine despite having access at the household level for different social, occupational and educational classes and religious groups. The survey was conducted amongst a representative sample of households selected randomly covering almost the entire geographical area of the country. A stratified two-stage design was adopted for the NSS 76th round survey where the first stage units (FSU) were villages/Urban Frame Survey (UFS) blocks/Sub-units (SUs) and the Ultimate Stage Units (USU) were households in both sectors. For the central sample, the total number of FSUs surveyed for Schedule 1.2 of NSS 76th round was 8,992 at the all-India level of which 5,378 was in rural areas and 3,614 was in urban areas. A sample of 63,736 rural households and 43,102 urban households was surveyed. The data used in this study is based on the unit-level data of this survey. Data for 'Use of Latrine' was Block 3 of this survey in which data is collected at the individual level. The total number of persons surveyed was 4,66,527 all over the country.

(B) Input Variables for Machine Learning Models

On the basis of intuition, common knowledge and ease of collection 16 variables are selected as input features: 'Gender', 'Age', 'Marital status', 'Highest level of education', 'Usual principal activity status', 'Sector', 'State_Code', 'Household size', 'Religion', 'Social group', 'Land possessed as on date of survey', 'UMCE' (Usual Monthly Consumer Expenditure), 'Tenurial status of dwelling', 'Whether drinking water sufficient', 'Principal source of water excluding drinking', 'Whether gets sufficient water'.

When a correlation analysis of the above variables is done, 'Age' is found to be highly correlated to 'Marital Status' and 'Whether drinking water sufficient' is decently correlated with 'Whether gets sufficient water'. Hence, 'Marital Status' and 'Whether drinking water sufficient' were removed from the list of input variables. The Heatmap of the correlation matrix is shown above and after dropping these two variables, a total of 14 input variables are taken to be our super set. From this super set, we will randomly create a few sets and make some models on those input sets to see how the model is performing.

Figure 1. Heatmap of the Correlation matrix



Note 1. In general, increasing the number of input variables increases the predictive ability of most models. But this is not always true. **It is much more important to use fewer 'quality' input variables rather than just increasing the number of input variables randomly.** Sometimes, blindly increasing the input variables leads to increased noise in the model which consequently decreases model performance.

Note 2. Randomly increasing the input variables is also not very user-friendly as it may be tedious to get the values of all the input variables to do predictions. For example, a model created on 40 input variables may perform fantastically on the test set. But, it might not be possible to get values of all 40 input variables every time to do predictions and thus, putting a big question mark on the usability of such bigger models in real life and therefore, simpler models are often preferred.

(C) Dependent Variable and its Simplification

'Use of Latrine' is a categorical variable with the following categories:

regularly 1

occasionally 2

never 3

In its present form, this is a multi-class classification problem because the dependent variable has more than 2 categories. For this paper, we will transform this into a binary classification problem with new codes using the logic explained below.

The total number of individuals surveyed in the NSS 76th round is 4,66,527. The distribution of 'Use of Latrine' is given below:

Table 1. Original Distribution of 'Use of Latrine (code)'

Use of Latrine (code)	Counts
1	3,73,865
3	9,956
2	2,873
Total	**3,86,694**

The total number of individuals where 'Use of Latrine (code)' is not available is 79,833 (4,66,527-3,86,694=79,833). According to the related instructions in section 3.3.11 of the instruction manual of this survey, columns 11 to 14 are relevant only for those households which have access to latrine i.e. for the households with any of the codes 1 to 4 and 9 in item 25 of block 5 of the schedule of enquiry. Entries in these columns are to be recorded after canvassing block 5.

It may be inferred from above that 79,833 people do not have access to latrines. Hence, it is assumed that these 79,833 people are not using latrines (code 3).

Further, code 2 which means occasionally using a latrine is also assumed as not using a latrine (code 3). There are two reasons for this assumption which are mentioned below.

i. The count of individuals with code 2 is extremely low (0.74%). So, reclassifying them as code 1 or code 3 or dropping them altogether from the analysis doesn't make a big difference.
ii. As per section 3.3.12.2 of the instruction manual, if any household member does not use latrine in most of the circumstances i.e. the common practice of the household member is not to use latrine but to go for open defecation, then it will be considered that the household member is not using latrine regularly and entry in this column will be recorded as either 2 or 3. Further, as clarified in the instructions, sometimes it may happen that though the common practice of the household member is to go for open defecation, the member may use latrine from time to time, say, in an emergency or during the rainy season only or in other circumstances. For such members of the household, code 2 is to be recorded. This is more suggestive of people not using latrines than using latrines.

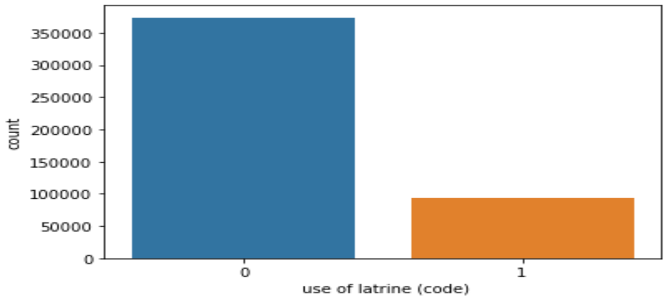
In the variable 'Use of Latrine', we are interested in the category of people who do not use latrines. So, this is our positive class and for modelling purposes will be re-coded as 1 (from earlier 3). People using latrines will be re-coded as 0 (from earlier 1).

The result of the above simplifications is the following binary distribution of the dependent variable:

Table 2. Simplified Distribution of 'Use of Latrine (code)'

Use of Latrine (code)	Counts
0	3,73,865
1	92,662

Figure 2. Count plots of Use of Latrine



From Table-2 and Figure-2 it is clear that the **dataset is imbalanced** with fewer positive classes (i.e. people not using Latrine).

(D) Preparing the Data for Machine Learning Algorithms

i) Dealing with Missing Values

Out of the total 14 input variables considered for this study only 'Highest level of education' data is missing for 46 persons. When data of these 46 individuals are analysed, it is found that the average age of this group is 7.04 years and the median age is 6 years. As the mean and median age are low, all these 46 missing values are replaced with 1 (not Literate).

ii) Splitting the Dataset

A total of 4,66,527 observations are present in the NSS 76th round dataset. We split this into a train set and test set in a 60:40 ratio. Generally, a split of 80:20 is done but to make a more robust model 60:40 split is done in this study. This means we will randomly use 2,79,916 (60%) data for training the model and 1,86,611 (40%) data for testing the model. The splitting is done in such a way that the percentage of '1' and '0' of the dependent variable remains the same in the train and test set as it was in the original data to avoid clustering of 1s or 0s in either of the sets.

3. Model Building

In this section, we shall build models using the training set and evaluate them on the test set. The machine learning algorithm used for model building is CatBoost Classifier. Four models with the same hyper parameters (learning rate 0.1, depth 10, scale-pos-weight 3, early-stopping-rounds 50) but a different set of input variables are trained and tested.

In the first step, three models: Model-1 comprising 7 variables, Model-2 comprising 10 variables and Model-3 comprising 14 variables (full set) are created. The input variables for these models are chosen heuristically from the superset of 14 variables (explained in section 2B).

In the second step, using the 'feature importance' method of CatBoost, we extracted the best 6 variables out of Model-3. These 6 variables are used as input to create Model-4. Only the best 6 features are chosen to create a model which is more modest (in terms of input variables) than the other three models created heuristically.

In other words, input variables of **Model-4 are NOT chosen heuristically but come from the machine learning part of Model-3**.

The confusion matrix, classification report, order of importance of input features and AUCPR (Area under the Precision-Recall curve) graph of each model is given below.

a) Model 1. 7 input variables ('Gender', 'Age', 'Highest level of education ', 'Sector', 'State_Code', 'Social group', 'UMCE').

Figure 3. Confusion Matrix of Model-1

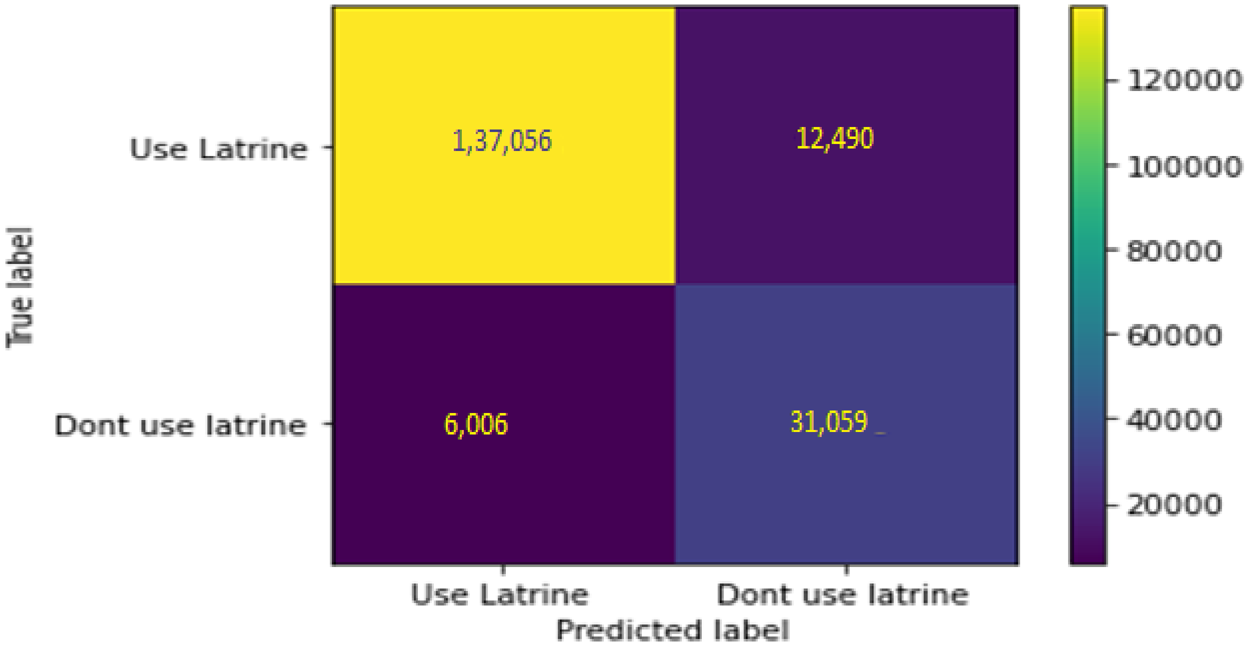


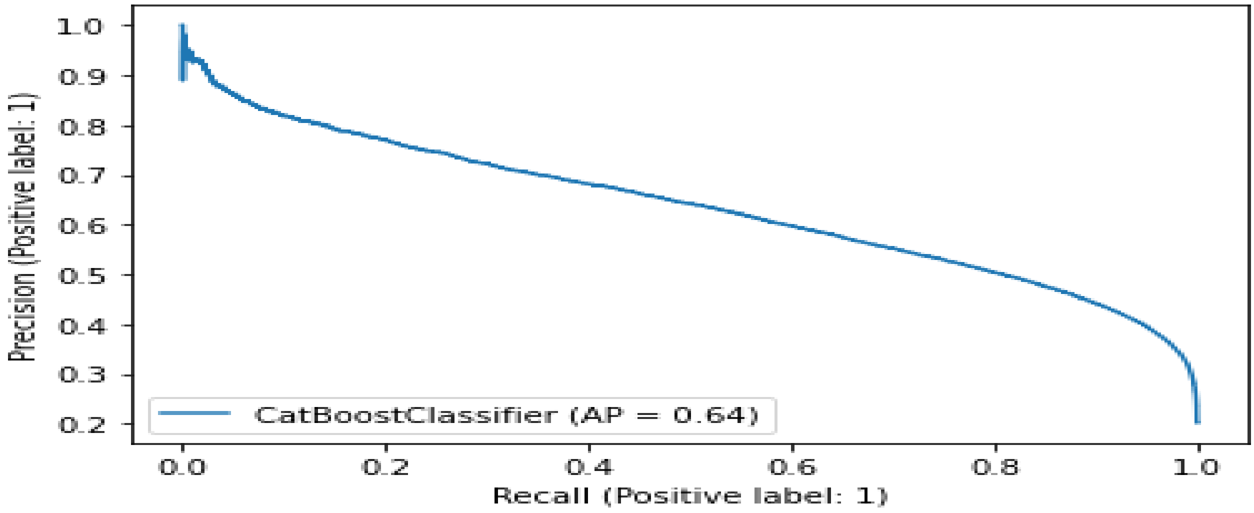
Table 3. Classification Report of Model-1

	Precision	Recall	f1-Score	Support
Use Latrine	0.94	0.81	0.87	1,49,546
Don't use Latrine	0.51	0.8	0.62	37,065
Accuracy	0.81	0.81	0.81	0.81
Macro Avg	0.72	0.8	0.74	1,86,611
Weighted Avg	0.85	0.81	0.82	1,86,611

Table 4. Order of Importance of Input Variables

Feature Name	Importance
State_Code	25.24
UMCE	20.31
Social Group	14.69
Sector	12.93
Age	11.36
Highest Level of Education	9.89
Gender	5.57

Figure 4. PR Curve of Model-1 (AUCPR = 0.64)



b) Model 2. 10 input variables ('Gender', 'Age', 'Highest level of education', 'Sector', 'State_Code', 'Household size', 'Social group', 'UMCE', 'Principal source of water excluding drinking', 'Whether gets sufficient water')

Figure 5. Confusion Matrix of Model-2

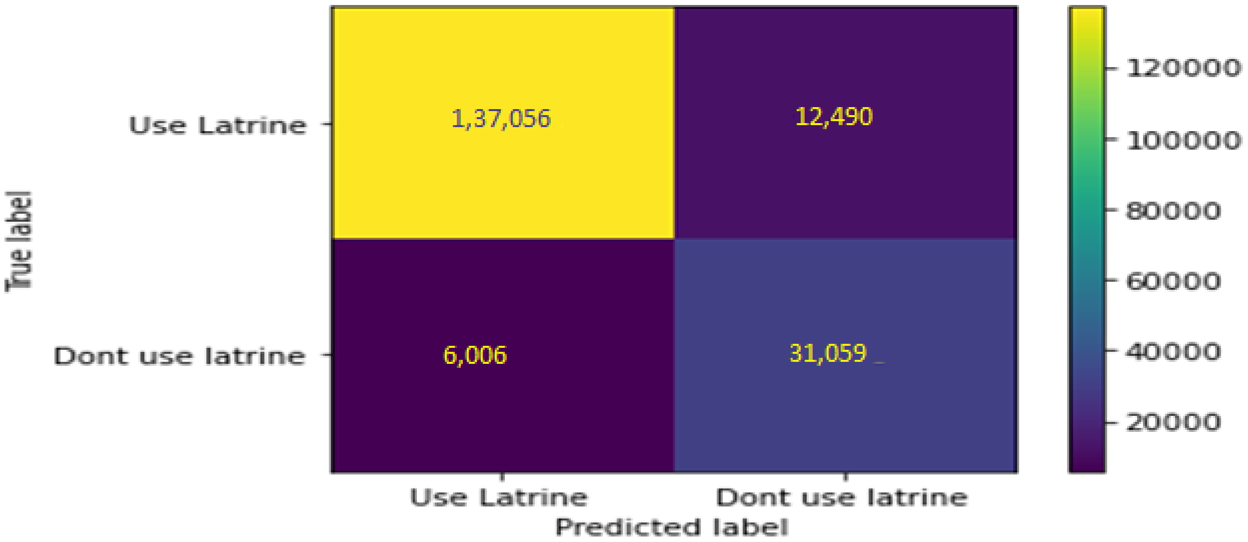


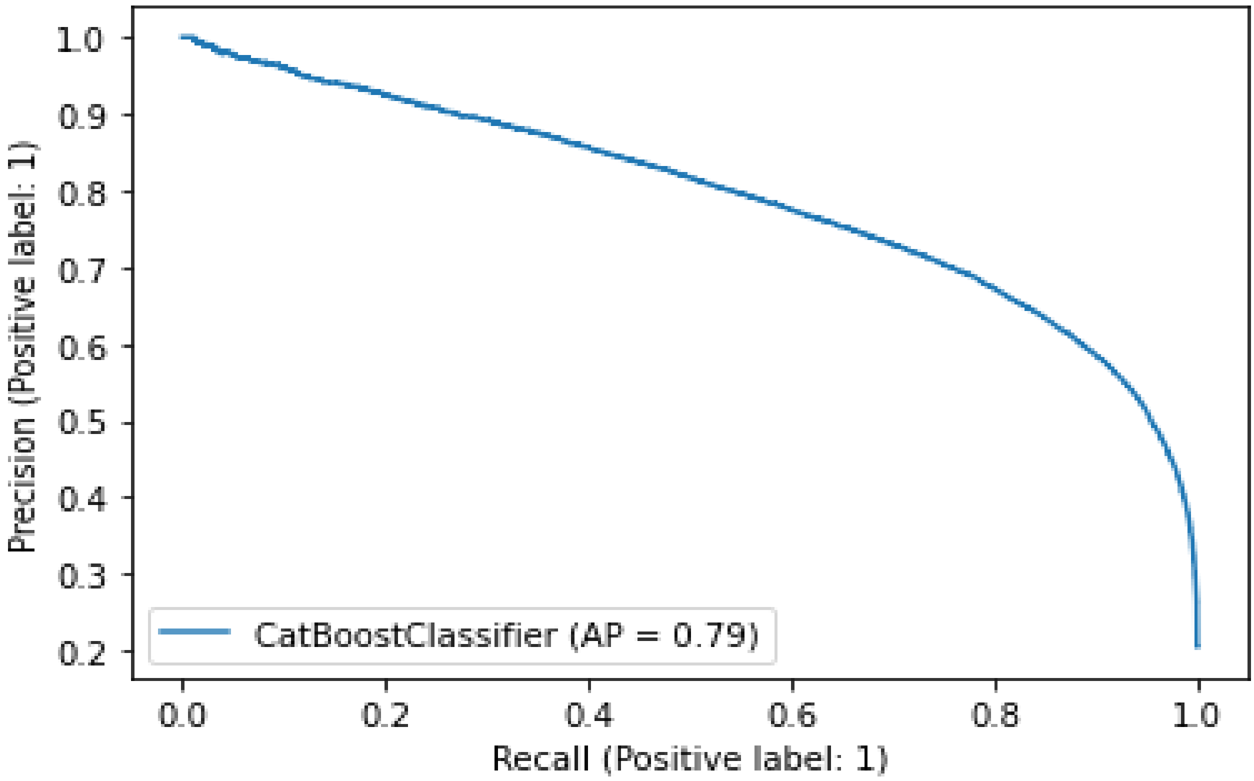
Table 5. Classification Report of Model-2.

	Precision	Recall	f1-score	Support
Use Latrine	0.95	0.89	0.92	1,49,546
Don't use Latrine	0.66	0.82	0.73	37,065
Accuracy	0.88	0.88	0.88	0.88
Macro Avg	0.8	0.86	0.83	18,6,611
Weighted Avg	0.89	0.88	0.88	18,6,611

Table 6. Order of Importance of Input Variables.

Feature Name	Importance
State_Code	20.13
UMCE	17.75
Principal Source of Water Excluding Drinking	12.60
Household size	11.72
Social group	11.06
Age	8.82
Highest Level of Education	6.69
Sector	6.01
Gender	3.57
Whether Gets Sufficient Water	1.66

Figure 6. PR Curve of Model-2 (AUCPR = 0.79)



c) Model 3. 14 input variables ('Gender', 'Age', 'Highest level of education ', 'Sector', 'Usual principal activity status', 'Religion', 'State_Code', 'Household size', 'Social group', 'Land possessed as on date of survey', 'UMCE', 'Tenurial status of dwelling', 'Principal source of water excluding drinking', 'Whether gets sufficient water').

Figure 7. Confusion Matrix of Model-3.

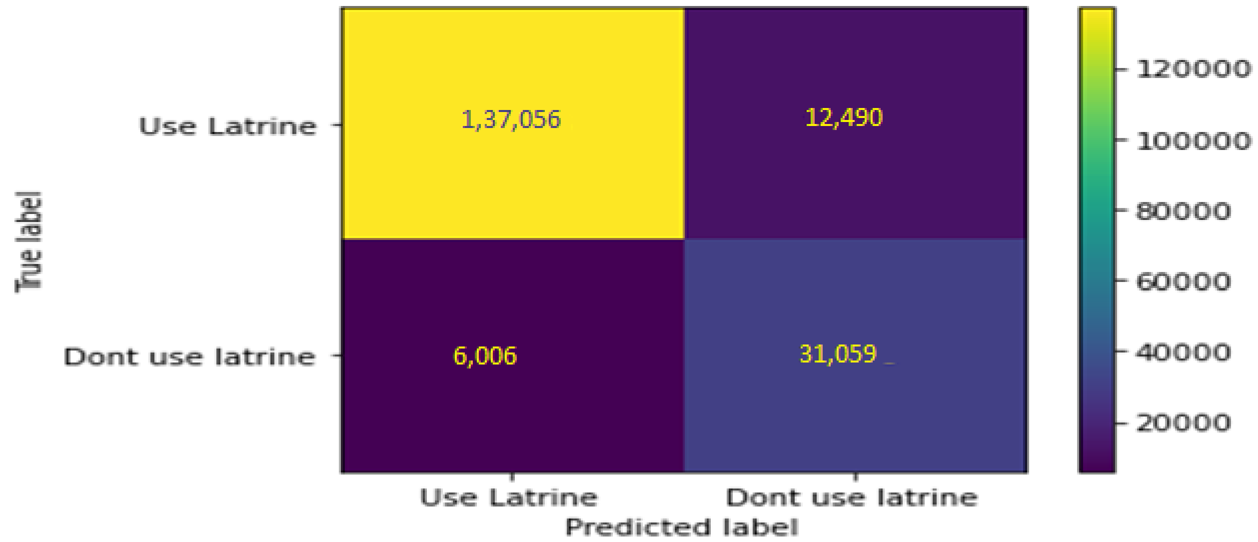


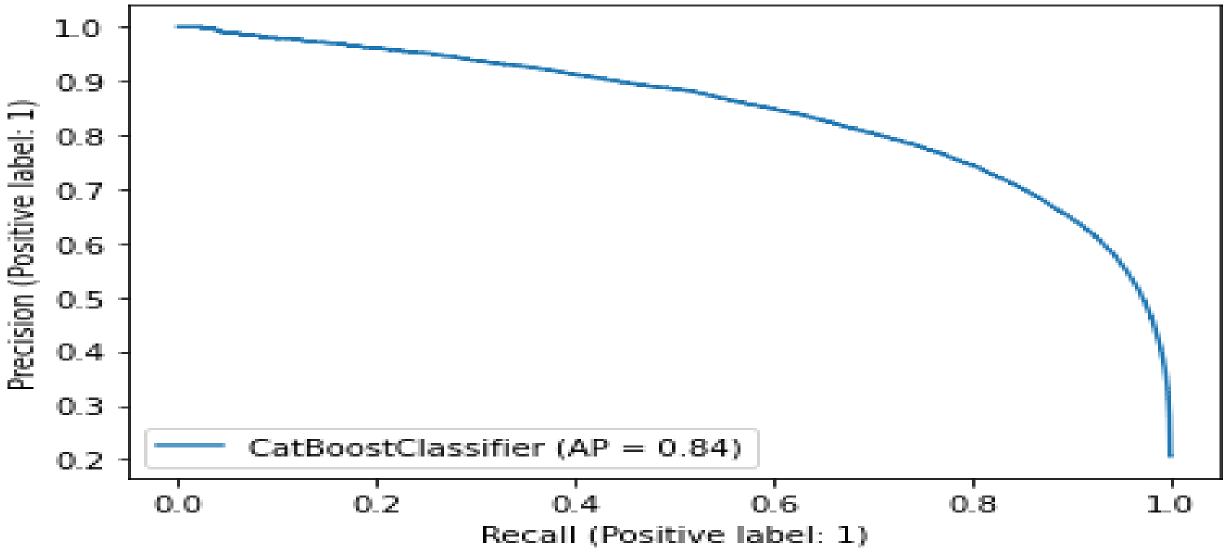
Table 7. Classification Report of Model-3.

	Precision	Recall	f1-score	Support
Use Latrine	0.96	0.92	0.94	1,49,546
Don't use Latrine	0.71	0.84	0.77	37,065
Accuracy	0.9	0.9	0.9	0.9
Macro Avg	0.84	0.88	0.85	1,86,611
Weighted Avg	0.91	0.9	0.9	1,86,611

Table 8. Order of Importance of Input Variables.

Feature Name	Importance
State_Code	18.13
UMCE	14.17
Land Possessed as on Date of Survey	10.03
Principal Source of Water Excluding Drinking	9.78
Social Group	9.06
Household Size	8.90
Age	6.54
Highest Level of Education	5.47
Sector	4.50
Usual Principal Activity Status	4.42
Religion	4.04
Gender	1.91
Tenurial Status of Dwelling	1.81
Whether Gets Sufficient Water	1.23

Figure 8. PR Curve of Model-3 (AUCPR = 0.84)



d) Model 4. Based on step c), we chose the top 6 variables out of the 14 ('State_Code', 'Household size', 'Social group', 'Land possessed as on date of survey[2]', 'UMCE', 'Principal source of water excluding drinking').

Figure 9. Confusion Matrix of Model-4

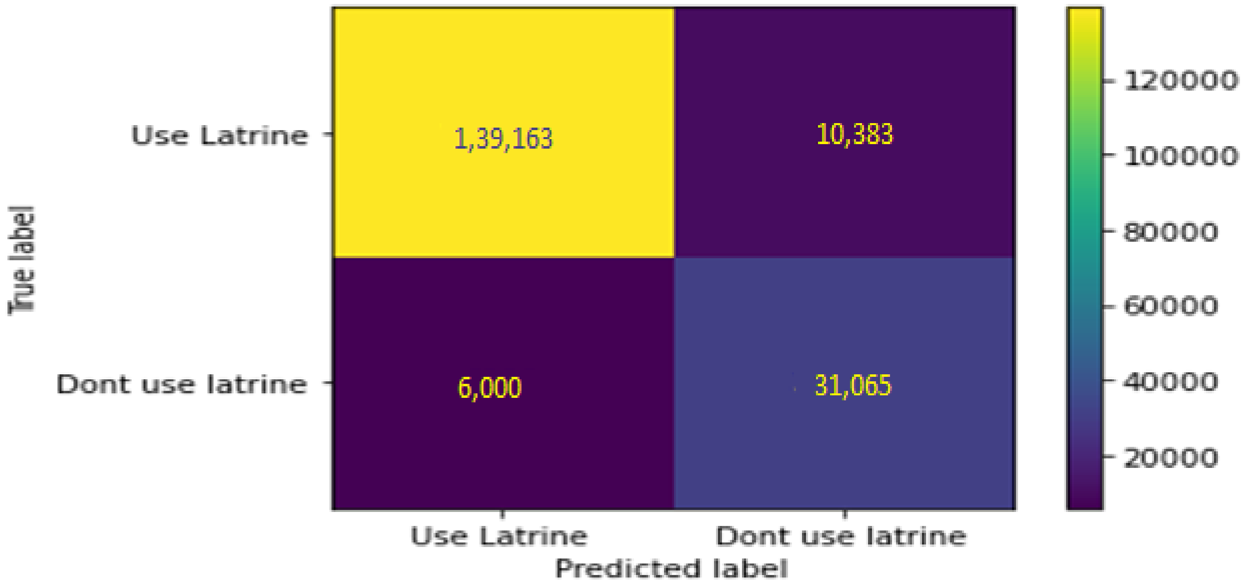


Table 9. Classification Report of Model-4.

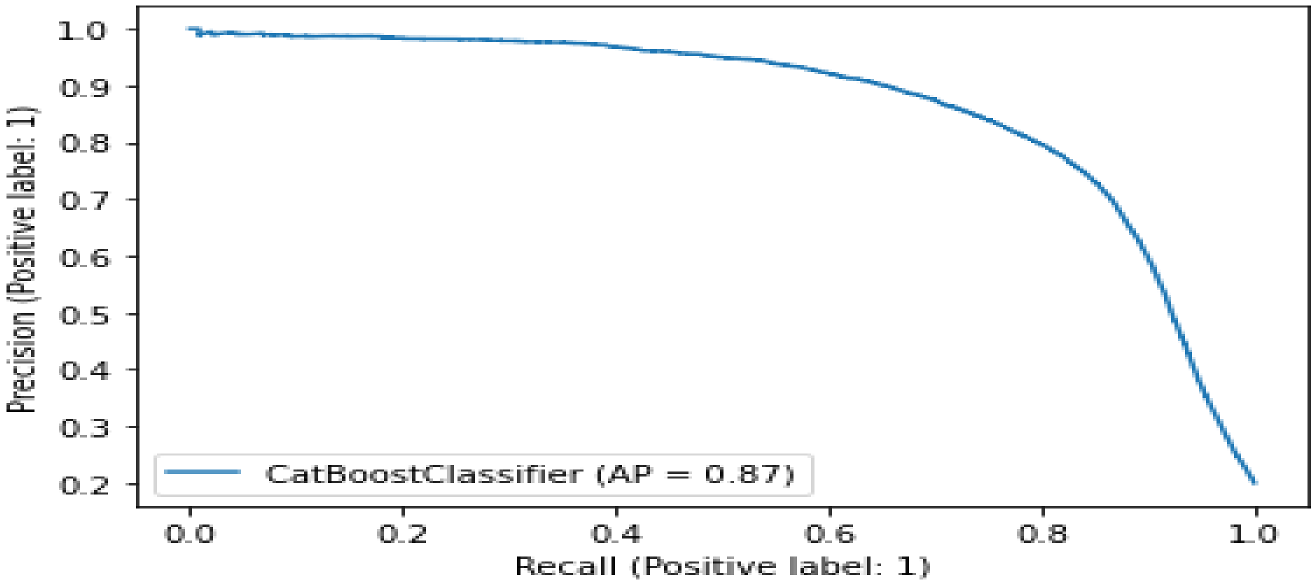
	Precision	Recall	f1-Score	Support
Use Latrine	0.96	0.93	0.94	1,49,546
Don't use Latrine	0.75	0.84	0.79	37,065
Accuracy	0.91	0.91	0.91	0.91
Macro Avg	0.85	0.88	0.87	1,86,611
Weighted Avg	0.92	0.91	0.91	1,86,611

[2] As regards the relevance of this variable for urban areas, it will be the same as it is for rural areas. When this variable (along with other variables) was used to construct Model 3, the performance of the model increased and as per the order of importance, this variable played a crucial role in improving model performance. Hence, this variable was used to create Model 4.

Table 10. Order of Importance of Input Variables.

Feature Name	Importance
UMCE	24.87
State_Code	20.77
Land Possessed as on Date of Survey	15.02
Household Size	14.08
Principal Source of Water Excluding Drinking	13.81
Social Group	11.45

Figure 10. PR Curve of Model-4 (AUCPR = 0.87)



4. Model Evaluation and Comparisons

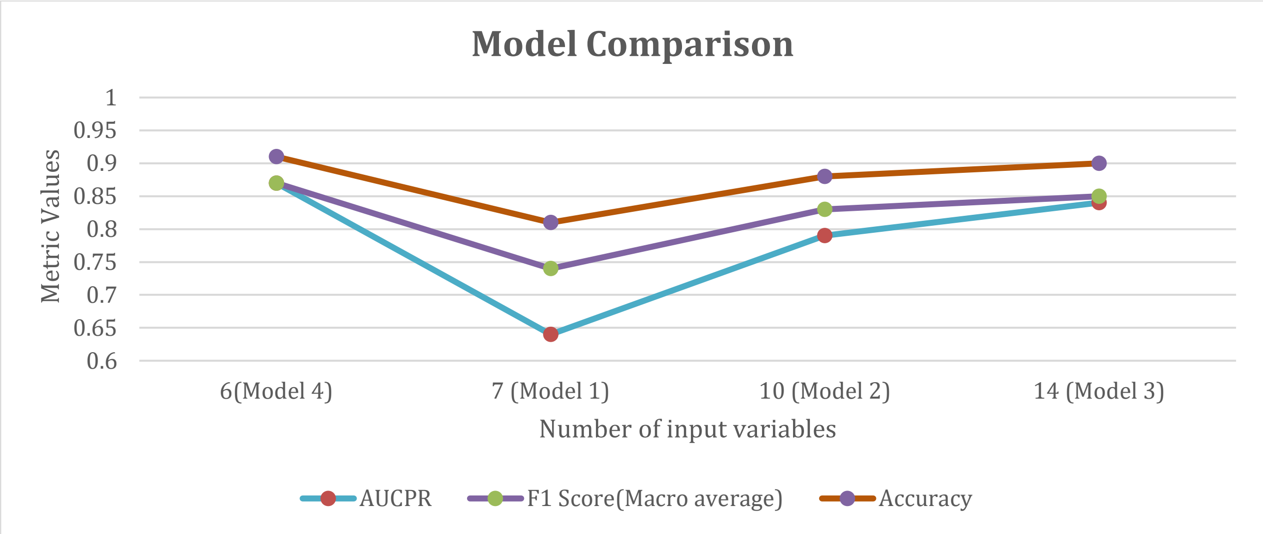
It is shown in section 2C that the dataset is imbalanced. Being an imbalanced dataset, accuracy and AUC will not be appropriate metrics to judge the performance of the models (Saito et al 2015). Therefore, in this case, we will be looking at the F1 score and AUCPR (area under the Precision-Recall Curve) for evaluating model performance.

Precision is a metric that quantifies the number of correct positive predictions made. It is calculated as the number of true positives divided by the total number of true positives and false positives. Recall is a metric that quantifies the number of correct positive predictions made out of all positive predictions that could have been made. It is calculated as the number of true positives divided by the total number of true positives and false negatives (e.g. it is the true positive rate). Both the precision and the recall are focused on the positive class (the minority class) and are unconcerned with the true negatives (the majority class).

F1 score is the harmonic mean of precision and recall. Precision-Recall curve is a function of precision(y) and recall(x). The baseline AUCPR is (total positive sample)/(total data samples). The baseline AUCPR for our data is 0.1986. The maximum value of AUCPR is 1. So, any value of AUCPR greater than 0.1986 will be considered an improvement.

The AUCPR of Model-1, Model-2, Model-3 and Model-4 are 0.64, 0.79, 0.84 and 0.87 respectively. Not only AUCPR but even accuracy, F1- scores of macro average are indicating the supremacy of Model-4 as shown in Figure 11. This implies that **Model-4 is the best among the 4 models even though it has the least number of input variables**.

Figure 11. Model Comparisons on Different Metrics



5. Conclusions

In this paper, we have been able to build a machine learning model with a recall value of 0.84 for the positive class. It means that the model can correctly identify 84% of people who do not use latrines in the population. The model uses only 6 input features and still performs better than other models using more input variables (7, 10, 14).

Due to fewer input variables, this model is simple, practical and handy. It can be used in real-life scenarios to identify individuals who do not use a latrine. Once identified, appropriate measures/schemes can be launched to encourage them to use latrines to accelerate the objective of 'Making India Open Defecation Free'.

It may also be noted that the input features of Model-4 are not based on intuition but is a subset of input variables used in other machine learning model (Model-3). In other words, the input variables of Model-4 are suggested by the machine itself. It shows that machine learning algorithms can not only build efficient models but can also be used to choose variables to build even better models. Though this research pertains to data for year 2018, we think that the model is innovative and could be applied to any latest data.

Without using machine learning we never would have thought that dropping critical variables (in the context of open defecation) like 'Sector' (Urban/Rural), 'Religion', 'Highest Level of Education', 'Age' etc. from a model could give better results. Model-4 indicates that 'State', 'UMCE' etc play a bigger role in determining latrine non-usage than 'Sector', 'Religion' etc. This is a very interesting and unique finding of this paper and opens a new door for further exploration in this direction.

Acknowledgment

I thank Sh. Abhimanyu Dixit for providing support in the machine learning part. I would also like to thank Ms. Chinneivah Haokip, ISS for her immense help in improving the paper. Lastly, I would like to thank anonymous referees and the Editorial Advisory Board (EAB) of this journal who provided valuable feedback at multiple levels for revisions in this paper.

References

1. Aggarwal, S. (2021). Is India Really Open-Defecation-Free? Here's What Numbers Say. (https://www.downtoearth.org.in/news/rural-water-and-sanitation/is-india-really-open-defecation-free-here-s-what-numbers-say-77918).
2. Government of India (2019). Key Indicators of Drinking Water, Sanitation, Hygiene and Housing Condition in India. National Sample Survey Office (NSSO), Ministry of Statistics and Programme Implementation, Government of India.
3. Saito, T. and Rehmsmeier, M. (2015). The Precision-Recall Plot is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE* 10: e0118432 (https://doi.org/10.1371/journal.pone.0118432).

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

Gurprit Grover[1], Radhika Magan[2]

Abstract

Quality-Adjusted-Life Years (QALY) is an important measurement of health outcome. Based on the technique of utility maximisation, discrete axiom of revealed preferences (DARP) and spline regression help us in the estimation of utility values. Based on these utility values, we estimate Quality Adjusted Life Year (QALY) for different states of India.

In this paper we have adopted an econometric approach of revealed preferences by ridge method and a statistical approach of fitting through spline regression for the computation of utility values. Data has been taken from the NSS 75th round 'Key indicators of social consumption in India: Health' (July 2017 – June 2018). The axiom approach not only leads to maximisation of utility but also the regression approach helps in formation of utility function based on expenditure incurred for in patients during the hospitalisation. This leads to a novel formulation of utility model which serve as a framework for computation of QALY.

QALY values obtained from DARP and spline method were consistent for few states while other states show a slight difference from both the methods. States whose QALY values are greater than 0.5 are considered in category A which indicates a better quality of life. The other states whose QALY values are greater than 0.25 but less than 0.5 are considered in category B which indicates a moderate quality of life. While the remaining states whose QALY values are greater than 0 but less than 0.25 are considered in category C which indicates a major improvement is required in order to better their quality of life.

This measure of health outcome will help the health economists in the allocation of health care resources for those states which have lower QALY values. New methods used in the estimation of functional forms of utilities leads to better estimation of QALY values and increase their accuracy.

Keywords: NSSO, DARP, Ridge Regression, Spline Regression, Utility, QALY.
JEL Code: C19, C65.
Receipt of Final Version of paper from Author: April 2023
Acceptance: April 2023

---

[1]Professor, Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi.
[2] Research Scholar, Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi.

1. Introduction

Economic evaluation is a systematic tool which evaluates the benefits derived from a health care technology against the cost incurred in its usage. In this paper, we consider a healthcare utilization study which has different perspectives as well as insights. For an individual or groups of individuals, earlier health was measured on the basis of presence or absence of a disease concern. Now days, measurement of health is described in a different way. It has moved to the amount of life lived and how far is the satisfaction level achieved after undergoing a treatment or surgical procedure for the disease concern. This has further led to the introduction of the concept of "Quality of Life". Harris (1985) discussed "Qualifying the value of life", by defining reasonable quality as well as extended quality of life. He describes QALY as a year of healthy life expectancy which is worth 1 year of perfect health. While a worth of less than one is a year of unhealthy life expectancy.

In a multiple regression model if there exist inter-correlations between two or more independent variables then the ordinary least square (OLS) technique fails in the dataset. Panay *et.al*. (2019) present an interpretable regression method based on the Dempster-Shafer theory in order to predict the health care costs using the Evidence Regression model. Thus, penalized linear regression gives a transparent picture for a multicollinear dataset. Kan *et.al*. (2019) compared the prediction performance of standard and penalized linear regression in predicting future health care costs in older adults.

To study the effect of a treatment in a particular state, there are different measures of health outcome. The main impact should be observed on the patient's length of life and health related quality of life. Thus, a good health outcome should capture both the impacts carefully. QALY's are defined and computed using different methods such as standard gamble, time trade off, person's trade off, visual method and multi-attributable scale. Apart from these techniques, in this paper we have estimated the coefficients of the dataset using ridge regression and then applying the DARP approach for computation of utility values. The second approach involves estimation of utility values by using spline regression. Both the techniques have been used for the first time in the estimation of QALY.

The concept of QALY which is being measured with the help of utility function is introduced in this paper. Depending on the past literature, there are very few techniques applied in the estimation of indirect utility function for the computation of quality of life. So, we aim to analyse and articulate two different methods used in different background setup but on the same dataset. In both the cases, utility is formed through indirect functions so to cross validate our results and interpretations we try to amalgamate two approaches on different grounds.

2. Utility

The utilitarian philosophers describe utility as a measure for increasing or decreasing the value for happiness. People desire for things or goods which in turn leads to maximization of positive utility (pleasure) or negative utility (pain). There is a controversy in defining the measurability of utility. Some authors define it as a measure of satisfaction which is subjective in nature or as an indicator of preferences which is objective in nature.

Nord (1994) proposed the usage of utilities based on QALY by means of societal values. QALY's are defined as the summation of utility adjusted values over various time intervals. There lies an underlying assumption for QALY to be of additive separability. It states that the computation of QALY is based on the underlying concept of utility of a given health state which is unaffected by the other health state that precedes or follow it. The usage of utilities in the formulation of QALY helps in decision making for healthcare when resources are scarce or limited.

Application of Afriat's theorem given in the paper by Grover *et.al.* (2019) indicates the presence of an increasing trend of utility function on a consumption space which rationalise the price and demand values. When goods are perfectly divisible then Afriat's (1967) theorem consider generalised axiom of revealed preferences (GARP) as the best condition for the evaluation of consumption data which becomes consistent with the maximisation of utility. Forges and Iehle (2014) consider DARP as a necessary and sufficient condition for rationalization of dataset. This further leads to the formulation of a good utility function. Under cost efficiency the utility function in a consumption space will follow the principles of GARP and DARP. However, in our case the consumption bundles are given as key indicators of health which are explained in an indivisible form.

Polisson and Quah (2013) states about the cost efficiency which will hold if the goods dealt by the modeller are being consumed in a continuous manner. They even pointed out on studying the consumer choice over a consumption space for indivisible goods, wherein DARP works as an elaborative version of GARP. By the minimalistic approach one can observe the consumer's budget which also reveals their preferences indirectly. Further in DARP's approach the maximisation of utility is being done with the help of ridge regression. The second approach deals with the computation techniques involved in spline regression.

3. QALY

Health outcome is defined as changes in health that result from a treatment measure or specific health care interventions. QALY is a metric used by health economists to evaluate new and innovative healthcare treatment for any particular disease. It is an important measurement of health outcome which gives the quality adjusted life years for an individual or groups of individuals. Drummond *et.al*. (1997) have introduced the quality of life which can be quantified by using the concept of utility. Various authors have combined the effects of health care interventions on mortality as well as on morbidity. Their definition of QALY goes around a single index which can be termed as common currency enabling comparison across different disease areas.

According to Drummond *et.al*. (2005), QALY is a summary measure which incorporates the impact on quantity as well as quality of life. However there lies a big concern on saving lives of patient's under treatment. Few authors have redefined this concern as, "Where should we spend whose money, to undertake what programs, to save whose lives and with what probability?" This question in turn raised by Zeckhauser and Shephard (1976) implies on how many lives are saved along with the justification of the resources expanded. Their definition defines health as a value weighted score defined with respect to a time period. This value is further measured in terms of preference, which is observed uniformly across all the individuals and is aggregated over them.

Recent research has stated QALY as a leading metric to perform cost effective analysis. Kurz and Gossen (2005) allows for standardized measurement of health outcomes which compare across different diseases and populations.

This index helps us to evaluate the quality of life achieved after undertaking a treatment or surgery upon the occurrence of a disease. The severity and recurrence of the disease creates a burden on a patient's health as well as expenditure incurred on it. Further, this index aims to evaluate on how far is the satisfaction level achieved after spending an amount on the treatment facilities.

4. Data

Data has been taken from the NSS 75st round 'Key indicators of social consumption in India: Health' (July 2017 – June 2018). The survey is conducted by National Sample Survey (NSS) organization, Ministry of Statistics and Program Implementation, Government of India, New Delhi. In this round detailed information was collected on profile of ailments, role of government and private healthcare facilities in providing healthcare, expenditure on medicines, medical consultation, investigation, hospitalisation etc which is shown in table 1 below. The survey period was from July 2017 till June 2018. The objective of the survey was to generate basic quantitative information about health sector all over the nation. The survey investigated the nature of ailments for which people of various ages were hospitalised, the extent of use of government hospitals, and the expenditure incurred on treatment received from government and private facilities. It collected data from 5,55,114 households spread over every district of the country. The rural households belonged to 8,077 randomly selected villages and the urban households to 6,181 randomly selected urban blocks. The total expenditure during the last 365 days for medical treatment was categorized under different parameters:

Package Component: This component comprises of packages of treatment involving specific surgical or non-surgical medical procedures such as Operation theatres (OT) charges, OT consumables, medicines, doctor's fees, bed charges etc. When treatment cost is available in the form of package with predetermined total cost then information for different constituents of the treatment is not separately available.

Doctor/Surgeon Fee: This is the total amount paid for doctor's or surgeon's fee. This fee is chargeable for the treatment imparted to the patient within the reference period for stay in the hospital.

Medicine: It accounts for the charge of medicine used during the treatment (including drips)

Diagnostic Test: This charges the fees of diagnostic test done for the patients within the reference period.

Table 1. Percentage Break up of Hospitalisation Expenses Incurred for Treatment during Stay at Hospital for Private Hospitals by States

State/UT	Package Component (%)	Doctor Fee (%)	Diagnostic Test (%)	Medicines (%)	Total Expenditure (INR)
Andhra Pradesh	23.6	23.3	13.2	21.8	8,190
Arunachal Pradesh	0.3	18.6	14.7	34.7	9,800
Assam	55.2	9.9	7.1	13.1	8,530
Bihar	44.4	10.4	8.9	21.0	8,470
Chhattisgarh	41.9	9.8	9.9	20.1	8,170
Delhi	87.4	3.5	2.3	2.5	9,570
Goa	74.3	8.0	2.7	6.2	9,120
Gujarat	35.4	18.2	9.4	18.5	8,150
Haryana	34.1	17.1	9.6	16.9	7,770
Himachal Pradesh	56.4	8.0	6.5	17.3	8,820
Jammu & Kashmir	59.7	7.6	7.0	16.6	9,090
Jharkhand	46.1	14.3	8.3	17.0	8,570
Karnataka	32.2	18.2	10.2	21.4	8,200
Kerala	16.7	16.8	12.1	23.6	6,920

State/UT	Package Component (%)	Doctor Fee (%)	Diagnostic Test (%)	Medicines (%)	Total Expenditure (INR)
Madhya Pradesh	28.6	14.5	11.8	22.7	7,760
Maharashtra	29.5	22.5	9.9	17.4	7,930
Manipur	67.4	9.3	2.9	11.5	9,110
Meghalaya	61.6	8.2	5.1	11.4	8,630
Mizoram	18.1	17.5	9.5	37.0	8,210
Nagaland	42.5	10.4	7.9	18.8	7,960
Odisha	31.6	22.7	9.0	23.4	8,670
Punjab	48.7	11.8	7.5	16.3	8,430
Rajasthan	45.1	9.7	8.5	21.7	8,500
Sikkim	34.9	22.4	9.2	11.3	7,780
Tamil Nadu	47.9	17.2	5.9	16.4	8,740
Telangana	38.5	25.7	8.6	13.7	8,650
Tripura	37.6	26.3	9.9	15.3	8,910
Uttarakhand	23.1	17.5	15.4	26.1	8,210
Uttar Pradesh	38.2	12.4	8.5	25.3	8,440
West Bengal	65.5	9.0	5.6	8.7	8,880
A & N islands	45.1	16.7	7.3	15.1	8,420
Chandigarh	50.6	11.5	8.4	14.6	8,510
Dadra & Nagar Haveli	17.7	38.3	7.1	12.8	7,590
Daman & Diu	0	42.4	17.5	17.9	8,780
Lakshadweep	28.6	10.1	9.7	16.0	6,440

Source: NSS, 75^{th} Round, India.

5. Methodology

5.1 Ridge Regression

The set of variables like package component, doctor's fee, medicines, diagnostic test were highly correlated with each other. The presence of collinearity was checked by correlation matrix given in table 2 below. The computations for all the methods have been done by using different packages from R software.

Table 2. Correlation Matrix for Explanatory Variables

Variables	PC	DF	M	DT
PC	1	-0.70	-0.70	-0.82
DF	-0.70	1	0.15	0.39
M	-0.70	0.15	1	0.65
DT	-0.82	0.39	0.65	1

From Table-2 we can observe that package component (PC) is highly correlated with doctor's fee (DF), medicines (M), and diagnostic test (DT). Diagnostic test is also moderately correlated with the medicines. Thus, OLS technique will give us biased estimates of the parameters and very large variances. So, we choose the technique of ridge regression which helps us to eliminate the effect due to collinearity.

Ridge regression is preferred in comparison to Lasso regression which reduces the estimated value of highly correlated variable to zero i.e. shrinkage of the coefficient occurs which further leads to the elimination of the variables from the model. We prefer to choose Ridge regression as we do not want to lose out any information related to the set of predictor variables. Using the glmnet package in R we can fit the generalised models by means of penalised likelihood method. It is based on the algorithm of cyclical coordinate descent which iteratively minimises the objective function for a given set of parameters when they further reach a convergence point.

By default, the glmnet package is taking the underlying distribution to be Gaussian. On choosing the family= "Gaussian" the function glmnet helps in standardization of the dependent variable to have a unit variance. It then unstandardized the coefficients before computing the lambda sequence. The parameter lambda λ is the tuning parameter which is chosen by cross validation. We have chosen $\lambda \geq 0$ and $\alpha = 0$ in order to perform ridge regression. This minimum value of λ is chosen by using the function lambda.min which gives the value at which the mean sum of squares is minimised.

The term shrinkage penalty is defined as lambda times the sum of squares of the coefficients which gets penalized when the coefficients become too large to handle. As the value of lambda increases, the shrinkage of the variables occurs which sometimes may lead to zero coefficients and help in further elimination as in the case of lasso regression. The parameter α helps to maintain balance between minimising residual sum of squares (RSS) defined as the amount of variance in the dataset which is not explained by the regression model itself and sum of squares due to coefficients. By using this regression technique, we can regularise the coefficients, improve the prediction accuracy and thus decrease the variance component.

Ridge equation: $NU = 85.52 - 0.00589PC - 0.0062DF - 0.0111DT - 0.00506M$ (1)

The above equation (1) is obtained by performing ridge regression. The dependent variable NU refers to the number of persons who are undertaking the treatment in a particular state. PC means the package component, DF refers to the doctor's fee, DT denotes the diagnostic test, Mare the expenditure incurred on medicines taken during hospitalisation.

5.2 Formulation of Utility Function

Method-I

DARP method. Consumption of a particular bundle of goods which are components of expenditure for an inpatient during hospitalisation with a defined budget is given by p.x, where *p* refers to total expenditure incurred in a particular state and x refers to the total number of persons who have been treated as an inpatient during hospitalisation.

Among all the states the average total expenditure will be the component used in the computation for r_t.

From Afriat's inequalities there exist $\varphi_1, \ldots\ldots, \varphi_{36}$and $\delta_1, \ldots\ldots, \delta_{36}$> 0 such that

$$\varphi_k \leq \varphi_j + \delta_j \gamma_{jk} \quad (2)$$

For all j, k, t= 1,....,36.

Let $u_j : N^k \to \mathcal{R}$ defined as $u(x) \leq \min(\varphi_t + \delta_t(p_t x_t - r_t))$ (3)

Where u(x) denotes the utility value for a particular state, δ_t refers to rationalization parameter which is required to define a utility function, p_t denotes the total expenditure incurred for an inpatient during hospitalisation in a particular state t, x_t refers to the number of patients who were registered for the treatment as in patient in a particular state, $r_t$is formed by multiplication of average value of expenditures incurred in a private hospital throughout the year with the number of patients in a particular state, i.e. $r_t = avg(p_t) * x_t$.

By using the above equation (3), the utility values have been obtained in the following table 3 given below:

Table 3. Estimation of DARP Utility Values for Different States of India

State / UT	In-Patient Count	Ψ(x)	u(x)
Andhra Pradesh	43	7.95	0.12
Arunachal Pradesh	35	2.41	0.41
Assam	17	14.16	0.26
Bihar	14	3.83	0.23
Chhattisgarh	28	7.20	0.13
Delhi	33	2.69	0.37
Goa	56	3.77	0.26
Gujarat	30	6.57	0.15
Haryana	31	2.07	0.48
Himachal Pradesh	36	5.80	0.17
Jammu & Kashmir	28	3.89	0.25
Jharkhand	30	13.31	0.07
Karnataka	28	8.38	0.11
Kerala	95	0.32	0.10
Madhya Pradesh	28	2.03	0.49
Maharashtra	33	3.10	0.32
Manipur	24	3.81	0.26
Meghalaya	16	9.95	0.10
Mizoram	30	8.85	0.11
Nagaland	16	3.37	0.29
Odisha	30	8.57	0.11
Punjab	31	8.75	0.41
Rajasthan	28	4.88	0.04
Sikkim	26	2.12	0.47
Tamil Nadu	34	6.96	0.14
Telangana	22	9.20	0.10
Tripura	45	4.94	0.52
Uttarakhand	23	8.85	0.11
Uttar Pradesh	31	6.66	0.15
West Bengal	47	5.19	0.19
A & N islands	52	1.98	0.50
Chandigarh	20	2.07	0.48
Dadra & Nagar Haveli	36	1.40	0.71
Daman & Diu	11	6.32	0.18
Lakshadweep	51	0.06	0.15
Puducherry	31	5.58	0.17

Method-II

Splines Regression. A higher degree polynomial may provide a good fit to the data but it creates the problem of over fitting. We can analyse such cases by means of residual sum of squares which becomes unstable in nature. Friedman (1991) presented a flexible regression modelling for high dimensional data which is based on technique of recursive partitioning. He further suggested on to fit an appropriate function in different ranges of explanatory variables known as piecewise polynomial. The joint points of such pieces are called knots and the polynomials obtained are known as splines. Thus, fitting of the model gets improved due to relaxed linearity assumptions under spline regression. Perperoglou (2019) explains the procedure for a regression model, in which splines are used to model the effects as a special case of multivariable regression, wherein some explanatory variables are non-linear in nature.

The polynomial fitting function in R fits the polynomial of order 12 which is higher order. It gives a biased estimate and leads to overfitting of the model. Then we performed spline regression. The B spline function in R software helps in curve fitting and numerical differentiation. This leads to the formation of marginal utility. B splines of order 7 are basis functions for utility formulation in our study. Prautzsch*et.al.* (2002) states that there is only one unique spline function which can be built as a linear combination of B splines.

By using Bsplines function in R we have fit the polynomial function for utility defined as:

$$u_t \sim f(Expenditure\ (TE), Number\ of\ persons(NU))$$

The functional form for u_t is obtained from equation (4) by differentiating with respect to number of persons and substituting the value expenditure by keeping it constant for a particular state. The utility function has been defined uniformly for all the states due to limited availability of data.

$$NU = 51.2 + 114.36(1-TE) - 78.83(1-TE)^2 - 5.17(1-TE)^3 - 37.97(1-TE)^4 \\ +12.53(1-TE)^5 - 35.29(1-TE)^6 - 15.02(1-TE)^7 \quad (4)$$

The value of goodness of fit for the above model is given as: R^2 = 68%. This implies that 68% of the of the total variation in number of persons who are treated as inpatients for private hospitals is explained by the total expenditure incurred in treatment for in patients during their stay in hospitals based on different states of India.

6. Result

The estimated utility values given in table 3 are monotonic in nature and satisfies the properties of a DARP function in presence of hospital unit as a consumption bundle for indivisible goods. These values further estimate the QALY denoted by Q_{DARP} . The utility values obtained in table 4 are indirect marginal utility values which further leads to the estimation of QALY denoted by$Q_{Bspline}$. On the basis of utility function and average length of stay in hospital (ALOS), QALY's for different states can be estimated by:

$$QALY = Utility\ * Averagelengthofstayinhospital \quad (5)$$

ALOS in the hospital has been defined below in table 5. This column is estimated by considering the average length of stay and inpatient count on an average in each state during the survey period. The estimated values of QALY's for different states of India are shown in Table-5 below:

Table 5: Estimated QALY Values for Different States of India

State / UT	**ALOS**	Q_{DARP}	$Q_{Bspline}$
Andhra Pradesh	3.58	0.15	0.28
Arunachal Pradesh	2.92	0.27	0.42
Assam	1.42	0.42	0.01
Bihar	1.17	0.18	0.41
Chhattisgarh	2.33	0.06	0.01
Delhi	2.75	0.28	0.32
Goa	4.67	0.50	0.40
Gujarat	2.50	0.15	0.02
Haryana	2.58	0.28	0.04
Himachal Pradesh	3.00	0.08	0.25
Jammu & Kashmir	2.33	0.30	0.28
Jharkhand	2.50	0.11	0.04
Karnataka	2.33	0.05	0.02
Kerala	7.92	0.92	0.84
Madhya Pradesh	2.33	0.46	0.45
Maharashtra	2.75	0.28	0.32
Manipur	2.00	0.42	0.26
Meghalaya	1.33	0.14	0.05
Mizoram	2.50	0.12	0.01
Nagaland	1.33	0.20	0.04
Odisha	2.50	0.04	0.01
Punjab	2.58	0.25	0.40
Rajasthan	2.33	0.38	0.38
Sikkim	2.17	0.25	0.30
Tamil Nadu	2.83	0.06	0.14
Telangana	1.83	0.17	0.50
Tripura	3.75	0.62	0.75
Uttarakhand	1.92	0.07	0.17
Uttar Pradesh	2.58	0.04	0.01
West Bengal	3.92	0.08	0.03
A & N islands	4.33	0.60	0.78
Chandigarh	1.67	0.05	0.03
Dadra & Nagar Haveli	3.00	0.04	0.67
Daman & Diu	0.92	0.19	0.17
Lakshadweep	4.25	0.01	0.04
Puducherry	2.58	0.48	0.28

Category A: This refers to those groups of states whose QALY values ≥ 0.5

Q_{DARP}&$Q_{Bspline}$: Kerala, Tripura, A&N islands, Dadra & Nagar Haveli.

Category B: This refers to those groups of states whose values $0.25 \leq QALY < 0.5$

Q_{DARP}: Arunachal Pradesh, Assam, Delhi, Goa, Haryana, Jammu & Kashmir, Madhya Pradesh, Maharashtra, Manipur, Nagaland, Punjab, Rajasthan, Sikkim, Chandigarh.

$Q_{Bspline}$: Andhra Pradesh, Arunachal Pradesh, Bihar, Delhi, Goa, Himachal Pradesh, Haryana, Jammu & Kashmir, Madhya Pradesh, Maharashtra, Manipur, Punjab, Rajasthan, Sikkim, Telangana, Puducherry.

Category C: This refers to those groups of states whose values $0 < QALY < 0.25$

Q_{DARP}: Andhra Pradesh, Bihar, Chhattisgarh, Gujarat, Himachal Pradesh, Gujarat, Jharkhand, Karnataka, Meghalaya, Mizoram, Odisha, Tamil Nadu, Telangana, Uttarakhand, Uttar Pradesh, West Bengal, Daman & Diu, Lakshadweep, Puducherry.

$Q_{Bspline}$: Assam, Chhattisgarh, Gujarat, Jharkhand, Karnataka, Meghalaya, Mizoram, Nagaland, Odisha, Tamil Nadu, Uttarakhand, Uttar Pradesh, West Bengal, Chandigarh, Daman & Diu, Lakshadweep.

7. Conclusion

An economic evaluation of a health care program can be done in different ways. It can be descriptive in nature based on burden of disease or cost of illness. This is a novel technique based on estimation of QALY values by using DARP approach. There is no previous literature based on computation of quality of life by using the axiom of revealed preferences. Spline regression has been introduced for the first time in estimation of quality of life from expenditure data from hospitals. Smoothing spline are powerful functions than fitting by higher order polynomials for estimating relationships between the variables.

QALY is termed as a cornerstone of economic analysis which combines morbidity gains and mortality impact of a treatment. To study the effect of a treatment on a particular state, there are different measures of health outcome. The main impact should be observed on the patient’s length of life and health related quality of life. Thus, a good health outcome should capture both the impacts carefully. The method of ridge regression does not help in variable selection as Lasso regression. While lasso regression on the other hand sets the variables to exact zero which ridge method does not do. In our present study we aim to reduce the effect of collinearity and not reduce the number of variables which could lead to variable selection and loss of information.

Echenique*et.al*. (2011) believed that an axiom like DARP based on revealed preferences helps in the formation of quasilinear utility function consisting of observed prices and bundles of continuous and discrete goods. This further rationalise the purpose of DARP methodology used in our study. As per the conservative version of Afriat’s theorem, Forges and Iehle (2013) states that a consumer behaves as a utility maximizer when we define a feasibility matrix associated with his revealed preferences choices as cyclically consistent. Thus, the data in our study consist of finitely many observed prices and consumption bundles for different states so that the consumer could afford depending on his ailment and budget.

The classification of different states under category A shows exact consistency from DARP and splines method. These states have better QALY values in comparison to other states. Thus, a little improvement in the allocation of health care resources by working on the variables which are stated as break up of expenditure can uplift these states to a condition of perfect health. Under category B, there are also moderate number of states which shows consistency with QALY values greater than 0.25 but less than 0.5. Exceptions to this case are 5 states like Andhra Pradesh, Bihar, Himachal Pradesh, Telangana and Puducherry have moderate QALY values by spline method but from DARP method they lie in category C. Under category C, there are few numbers of states which shows consistency with QALY values greater than 0 but less than 0.25. Exceptions to this case are 3 states like Assam, Nagaland, Chandigarh has low QALY values by spline method but from DARP method they lie in category B. Thus,a little difference in the number of states which are less consistent from both the methods in category B and C needs greater attention in order to improve their quality of life.

Acknowledgement

We would like to thank NSSO, MoSPI for providing the data based on key indicators of health.

References

1. Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International economic review*, *8*(1), 67-77.
2. Drummond, M.F., Stoddart, G. L., Torrance, G.W., O'Brien. B. (1997). Methods for the economic evaluation of health care programmes. *Oxford Medical Publication*, 2nd edition; 305.
3. Drummond, M.F., Sculphe, M.J., Torranu, G.W, O'Brien, B.J., Stoddart, G.L. (2005). Methods for the economic evaluation of health care programmes. *Oxford University Press*.; 379.
4. Echenique, F., Lee, S., & Shum, M. (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy*, 119(6), 1201-1223.
5. Friedman, J. H. (1991). Multivariate adaptive regression splines. *The annals of statistics*, 1-67.
6. Forges, F., and Iehlé, V. (2013). Essential data, budget sets and rationalization. *Economic Theory*, 54(3), 449-461.
7. Forges, F., and Iehlé, V. (2014). Afriat's theorem for indivisible goods. *Journal of Mathematical Economics*, 54, 1-6.
8. Grover, G., Das, R. N., & Magan, R. (2019). On the estimation of QALD (Quality Adjusted Life Days) based on expenditure data for Childbirth and Maternity services in India using utility theory. *Indian Journal of Economics and Development*, *7*(6), 1-9.
9. Harris, J. (1985). Qualifying the value of Life. *Journal of Medical Ethics*. 13, 117-123.
10. Kan, H. J., Kharrazi, H., Chang, H. Y., Bodycombe, D., Lemke, K., & Weiner, J. P. (2019). Exploring the use of machine learning for risk adjustment: A comparison of standard and penalized linear regression models in predicting health care costs in older adults. *PloS one*, 14(3), e0213258.
11. Kurz, H.D. and Gossen, H.H. (2016). Handbook on the history of economic analysis.*Edward Elgar Publishing*. 816.
12. National Sample Survey Office. (2019). Key Indicators of Social Consumption in India: Health.
13. Nord, E. (1994). The Qaly – a measure of social value rather than individual utility?*Health Economics*. 3(2), 89-93.
14. Panay, B., Baloian, N., Pino, J. A., Peñafiel, S., Sanson, H., &Bersano, N. (2019). Predicting Health Care Costs Using Evidence Regression. In *Multidisciplinary Digital Publishing Institute Proceedings* .31(1), 74.
15. Perperoglou, A., Sauerbrei, W., Abrahamowicz, M., & Schmid, M. (2019). A review of spline function procedures in R. *BMC medical research methodology*, 19(1), 1-16.
16. Prautzsch, H., Boehm, W., & Paluszny, M. (2002). *Bézier and B-spline techniques*, Berlin: Springer. 6.
17. Polisson, M., & Quah, J. K. H. (2013). Revealed preference in a discrete consumption space. *American Economic Journal: Microeconomics*, *5*(1), 28-34.
18. Bruce, P., & Bruce, A. (2017). Practical statistics for data scientists, OREILLY Sebastopol, California-USA.
19. James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. (2014). An Introduction to Statistical Learning: With Applications in R. Springer Publishing Company.
20. Zeckhauser, R. and Shepard, D. (1976). Where now for saving lives? *Law and Contemporary Problems*. 5-45.

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

P. D. Joshi[1] and Jena P Joshi[2]

Abstract

This paper focuses on (a) gender-specific mobility measures of persons in Current Weekly Activity Status (CWS) on quarterly basis and (b) inter-temporal quarterly changes in the urban areas of India. Measures have been computed using both Joshi and Singh's (1977) Mobility Measure (*D*) under the homogeneous Markov chain model based on entropy and Joshi's (2021) measure (*J*) of dynamic mobility for changes in activity status of persons based on Markov chain separately for males and females.

Data available from the first quarterly bulletin on Periodical Labour Force Survey (PLFS), started in 2017 by National Sample Survey Office (NSSO) under the Ministry of Statistics and Programme Implementation, Government of India, corresponding to the first quarter ending December 2018, to the fourteenth quarterly bulletin corresponding to the quarter ending December 2021 have been used. The dataset covers the first phase of the nationally enforced lockdown for COVID-19 outbreak as a national crisis towards the middle of March 2020 and the impact of nationwide sudden lockdown from 24th March 2020 causing notably atypical mobility of persons in activity status. The first phase of lockdown was in quarter ending March 2020 to quarter ending June 2020. The field work of PLFS was suspended from 19th April 2021 in some areas due to second wave of COVID-19. The lockdown was gradually lifted only from the middle of May 2021 onwards and gradually normalcy restored since the first week of June 2021. It affected livelihoods and employment status of persons across the country. Thus, our measures of mobility in activity status (*D* and *J*) before and after the pandemic periods, reflect disruptions in economic activity on account of the then livelihood and employment status of persons in the urban areas of India.

Keywords: Activity status, Labour Force, Dynamics, Mobility, Measure, Transition, Markov Chain
JEL Code: I16
Receipt of Final Version of paper from Author: April 2023
Acceptance: April 2023

---

[1]Former Deputy Director General, NSSO, Government of India. E-mail: joshipd@hotmail.com.
[2] Co-author is Associate Professor, Dr. D. Y. Patil Institute of Management and Research, Pune.
The views expressed are personal and not of the organization to which the authors belong.

1. Introduction

National Sample Survey Office (NSSO), Ministry of Statistics and Programme Implementation, Government of India, has been conducting Periodical Labour Force Surveys (PLFS) since April 2017 on quarterly basis in urban areas and annual surveys in rural areas of India. So far five annual reports of PLFS corresponding to the periods July 2017 to June 2018, July 2018 to June 2019, July 2019 to June 2020, July 2020 to June 2021 and July 2021 to June 2022 covering both rural and urban areas giving estimates of all the important parameters of employment and un-employment have been released. Besides the annual reports, seventeen quarterly bulletins of labour force indicators in the urban areas of India from first quarterly bulletin corresponding to the quarter ending December 2018 to seventeenth quarterly bulletin corresponding to quarter ending December 2022 of Periodical Labour Force Surveys (PLFS) have been released. The objective of the survey is to provide labour force data at a more frequent time interval in the form of quarterly bulletin giving labour force indicators. Prior to this, NSSO used to conduct quinquennial large-scale sample surveys on employment and un-employment based on stabilised concepts and definitions given in Annex. Findings of these surveys were used for planning, policy formulation and decision support as well as input for further statistical exercises by various Government organisations at the national and state levels, academicians, researchers and scholars. However, PLFS provides level and change in point estimate of the key labour force indicators viz. Labour Force Participation Rate (LFPR), Worker Population Ratio (WPR), Un-employment Rate (UR) in a short time interval of three months separately for males and females of age 15 years and above in Current Weekly Status (CWS). These indicators are subject to fluctuation of sampling and do not provide statistical dimension of mobility (movement) in activity status of persons in current weekly status.

2. Objectives

The objective of this paper is to measure dimensions of dynamic mobility for gender-specific differences, and inter-temporal differences in short periods i.e. on quarterly basis based on suitably defined measure of mobility resting on the Markov Chain. Efforts have been made to present the dynamics of the labour force based on data analysis. Broad conclusions have been drawn on the impact of COVID-19 outbreak that appeared as a national crisis towards the middle of March 2020, especially on the impact of nationwide sudden lockdown from 24th March 2020 on measures of mobility of persons in activity status. The first phase of lockdown was in the quarter January-March 2020 and the quarter April-June 2020. Thus, our analysis provides results on mobility measures before and after the pandemic in regard to disruptions in economic activity. The paper is therefore more informative in nature and conclusions may be meaningful and useful to the users of labour force data.

3. Existing Literature on Dynamic Mobility

Dynamic mobility refers to the movement of individuals from positions possessing a certain rank to positions either higher or lower in the social system. In this context, several scholars including Prais (1955), Matrass (1950), Bartholomew (1967), Joshi and Singh (1977), Mukherjee and Basu (1979), Mukherjee and Chattopadhyay (1956) have suggested measures of economic and social mobility for representing transitions over generations and over time. However, measure of dynamic mobility for changes in economic activity status categorised into three groups viz., employed, un-employed and out of labour force has not received any attention except Joshi and Singh's (1977) measure of mobility under the homogeneous Markov Chain model using entropies i.e. independency of the class structure. It is

i) Well defined continuous function of elements $p_{ij}{}'$s of the transition matrix P,
ii) Independent of the ordering of the classes,
iii) Finite and attains absolute minimum when there is no mobility,
iv) Monotonic, non-decreasing under some realistic conditions and attains its maximum when and only when the system attains some ideal situation.

Joshi and Singh's (1977) measure of mobility is given by $D = -\sum_i \sum_j p_{ij} Log\left(p_{ij}\right)$, where p_{ij} is the probability of moving (transition) from activity state ''i'' (i = 1, 2, 3) in a time period to activity state ''j'' (j = 1, 2, 3) in another time period. However, this measure is not suited if one or more elements in transition probability matrix (P) is zero. Recently, Joshi (2021) has derived a measure of dynamic mobility (J) based on Markov Chain by taking the position of complete *immobility* as the point of comparison. It weights the value of the off-diagonal elements by their distance from the diagonal. Accordingly, the measure is

$$J = \frac{\delta}{\left[\frac{1}{k-1}\sum_i \sum_j (i-j)^2\right]}$$

Where, $\delta = \sum_i \sum_j (i-j)^2 p_{ij}$ and the transition probability matrix $P = \left[p_{ij}\right]_{k \times k}$ and k = 3 (here) i.e. number of activity statuses in the past period of time (i = 1, 2, 3) and number of activity statuses in the current period of time (j = 1, 2, 3). The numbers 1, 2, 3 stand respectively for three activity statuses, viz. ''Employed'', ''Un-employed'' and ''Out of Labour Force''. Thus, the measure (J) is free from the short comings of Joshi and Singh's mobility measure (1977). Further, the relevance of the measure J in our context has already been stated in our earlier paper in Sarvekshana, the journal of NSSO, vol.110 & 111 (combined). The table gives the distribution of the persons who were surveyed in both the adjacent quarters.

In this paper, ''Employed'', ''Un-employed'' and ''Out of Labour Force'' are the three activity statuses, i.e. three states in Markov's terminology, for a person in a time span of current time period and time span of past (adjacent) time period. We have persons in either of these three activity statuses (states) in current period "j" and the same three activity statuses (states) ''i'' in the adjacent past period.

4. Data Base

The paper is wholly based on secondary data on employment and un-employment in urban areas of India drawn from the quarterly bulletins of PLFS. It provides a bivariate table on percentage distribution of persons of age 15 years and above in different activity statuses based on their current weekly status separately for males and females in adjacent quarters. A table was therefore compiled to have mobility in activity status of persons (males and females) categorised into three groups viz., "Employed", "Un-employed" and "Out of labour force" from a list given in Annex. The procedure for computing $p_{ij}{}'$s using the published data in the bulletins is as under.

A note in the bottom of Table-5 in quarterly bulletins mentions that the table gives the distribution of the persons who were surveyed in both the adjacent quarters. Further, Table-1 in quarterly bulletins provides the number of persons surveyed by age and gender in the age group 15 years and above. Using data from these two tables, the number of persons in i^{th} activity status in a past adjacent survey period $\left(N_{i\bullet}\right)$ and number of persons in j^{th} activity status in a current (adjacent) survey period $\left(N_{\bullet j}\right)$ such that $\sum_i N_{i\bullet} = \sum_j N_{\bullet j} = N$ (total number of persons). With this, the number of persons in i^{th} activity status in past time period as well as in j^{th} activity status in current time period $\left(N_{ij}\right)$ has been worked out for urban areas at the all-India level.

5. Sample Design

In PLFS, the sampling design adopted is a stratified multi-stage design. The First Stage Units (FSU) are the Urban Frame Survey (UFS) blocks in urban areas. The Ultimate Stage Units (USU) are households. A rotating panel design has been used in urban areas. In this rotational panel scheme, each selected household in urban areas is visited four times – first time with first visit schedule and another three times with revisit schedule. Information is collected for all members of the sample households using Computer-Assisted Personal Interviewing (CAPI) method with inbuilt validation rules on tablets (hand-held electronic devices). Sample sizes of USUs in Quarterly Surveys are given in Table-1. The fieldwork during COVID-19 pandemic was suspended from the middle of March 2020 and resumed in June 2020. During this period around 79% data was collected over telephone.

Table 1. Sample Size of Households and Persons in Quarterly Surveys.

Survey Period	FSU (UFS Blocks)	Household	Male	Female	Person
April-June 2018	5,739	44,697	91,561	89,219	1,80,808
July-Sept 2018	5,745	44,887	90,889	88,276	1,79,193
Oct-Dec 2018	5,743	44,963	90,403	87,537	1,77,966
Jan-March 2019	5,740	45,024	90,469	87,167	1,77,660
April-June 2019	5,723	45,288	91,522	87,852	1,79,422
July-Sept 2019	5,720	44,471	69,547	68,550	1,38,130
Oct-Dec 2019	5,722	45,555	71,085	69,788	1,40,906
Jan-March 2020	5,651	43,971	68,653	67,391	1,36,002
April-June 2020	5,635	43,209	67,887	67,084	1,34,990
July-Sept 2020	5,581	43,257	67,897	67,008	1,34,935
Oct-Dec 2020	5,563	43,693	68,349	67,420	1,35,801
Jan-March 2021	5,601	44,000	68,704	67,780	1,36,523
April-June 2021	5,619	43,892	68,047	66,978	1,35,060
July-Sept 2021	5,676	44,272	68,453	67,263	1,35,740
Oct-Dec 2021	5,697	44,533	69,027	67,590	1,36,636

Note: Male + Female may not be equal to Person, as the person also includes Third Gender.

6. Methodology

For Markov chain, sum of transition probabilities p_{ij}'s (i = 1, 2, 3; j = 1, 2, 3) in each row equals one. It forms a chain as defined by Markov (1907) on the ground that it is a chance process and may be related to the outcome of a given experiment affecting the outcome of the next experiment. Therefore, the transition probabilities presented in Tables 3, 4 and 5 for persons, males, and females in the paper have been derived from Table-5 of the quarterly bulletins which provides activity status of persons of age 15 years and above, separately for males and females in percentage terms. i.e. (N_{ij}/N), ($N_{i.}/ N$) and ($N_{.j}/ N$) where N is the number of persons surveyed in both the adjacent quarters, N_{ij} equals number of persons in activity status i and in activity status j, $N_{i.}$ equals total number of persons in activity status i and $N_{.j}$ equals total number of persons in activity status j. Elements p_{ij}'s of the transition probability matrix P are the ratios $N_{ij}/N_{i.}$ = (N_{ij}/N)*($N/N_{i.}$). The sum of the probabilities in each row is one. Symbolically, the transition probability matrix for selected activity status is as under:

Table 2. Transition Probability Matrix (*P*)

<table>
<tr><td rowspan="6">P =</td><td colspan="3" rowspan="3">States i in Adjacent Past Period</td><td colspan="3">States j in Current Period</td><td rowspan="7">Retention and Gain (From)</td></tr>
<tr><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Employed</td><td>Un-employed</td><td>Out of Labour Force</td></tr>
<tr><td>S1</td><td>Employed</td><td>1</td><td>p_{11}</td><td>p_{12}</td><td>p_{13}</td></tr>
<tr><td>S2</td><td>Un-employed</td><td>2</td><td>p_{21}</td><td>p_{22}</td><td>p_{23}</td></tr>
<tr><td>S3</td><td>Out of Labour Force</td><td>3</td><td>p_{31}</td><td>p_{32}</td><td>p_{33}</td></tr>
<tr><td colspan="7">Retention and Loss (To)</td></tr>
</table>

7. Periodical Transition Probabilities

Tables 3, 4 and 5 present the transition probabilities p_{ij}'s of respectively persons, males and females aged 15 years and above in Current Weekly Status (CWS) by broad activity (employed, un-employed and out of labour force) status for selected quarter in the urban areas of India based on PLFS.

Table 3. Transition Probabilities p_{ij}'s of Persons Aged 15 Years and Above in CWS during Selected Quarters in the Urban Areas of India.

Period	**July-Sept 18**	**Oct-Dec 18**	**Jan-March 19**	**April-June 19**	**July-Sept 19**	**Oct-Dec 19**	**Jan-March 20**
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p_{11}	0.9476	0.9504	0.9525	0.9504	0.95529	0.9704	0.9461
p_{12}	0.0190	0.0165	0.0166	0.0165	0.01412	0.0091	0.0247
p_{13}	0.0333	0.0331	0.0309	0.0331	0.03059	0.0205	0.0292
p_{21}	0.2093	0.1628	0.2000	0.1628	0.19512	0.1579	0.1316
p_{22}	0.6744	0.6977	0.6667	0.6977	0.68293	0.7632	0.7632
p_{23}	0.1163	0.1395	0.1333	0.1395	0.12195	0.0789	0.1053
p_{31}	0.0242	0.0206	0.0225	0.0206	0.02996	0.0153	0.0174
p_{32}	0.0130	0.0075	0.0094	0.0075	0.00749	0.0057	0.0058
p_{33}	0.9628	0.9719	0.9682	0.9719	0.96255	0.9790	0.9768

Table 3 (Contd.). Transition Probabilities p_{ij}'s of Persons Aged 15 Years and Above in CWS during Selected Quarters for Persons in the Urban Areas of India.

Period	April-June 20	July-Sept 20	Oct-Dec 20	Jan-March 21	April-June 21	July-Sept 21	Oct-Dec 21
(1)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
p_{11}	0.7982	0.9620	0.9685	0.9559	0.9108	0.9609	0.9577
p_{12}	0.1338	0.0163	0.0121	0.0196	0.0526	0.0147	0.0141
p_{13}	0.0680	0.0217	0.0194	0.0245	0.0366	0.0244	0.0282
p_{21}	0.0698	0.3478	0.2295	0.4000	0.1190	0.2931	0.1778
p_{22}	0.8605	0.5978	0.7049	0.5538	0.8095	0.6379	0.7333
p_{23}	0.0698	0.0543	0.0656	0.0462	0.0714	0.0690	0.0889
p_{31}	0.0097	0.0315	0.0133	0.0133	0.0096	0.0169	0.0170
p_{32}	0.0058	0.0074	0.0038	0.0038	0.0058	0.0056	0.0057
p_{33}	0.9845	0.9611	0.9829	0.9829	0.9846	0.9775	0.9773

Table 4. Transition Probabilities p_{ij}'s of Males Aged 15 Years and Above in CWS during Selected Quarters in the Urban Areas of India.

Period	July-Sept 18	Oct-Dec 18	Jan-March 19	April-June 19	July-Sept 19	Oct-Dec 19	Jan - March 20
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p_{11}	0.9612	0.9643	0.9685	0.9673	0.9688	0.9362	0.9620
p_{12}	0.0239	0.0223	0.0121	0.0164	0.0148	0.0053	0.0234
p_{13}	0.0149	0.0134	0.0135	0.0164	0.0163	0.0585	0.0146
p_{21}	0.2540	0.2222	0.2500	0.2031	0.2167	0.1000	0.1481
p_{22}	0.6508	0.6825	0.6618	0.7031	0.7000	0.8000	0.7778
p_{23}	0.0952	0.0952	0.0882	0.0938	0.0833	0.1000	0.0741
p_{31}	0.0337	0.0303	0.0301	0.0303	0.0414	0.0139	0.0230
p_{32}	0.0262	0.0265	0.0051	0.0152	0.0188	0.0038	0.0153
p_{33}	0.9401	0.9432	0.9511	0.9545	0.9398	0.9823	0.9617

Table 4 (Contd.). Transition Probabilities p_{ij}'s of Males Aged 15 Years and Above in CWS during Selected Quarters in the Urban Sector of India.

Period	April-June 20	July-Sept 20	Oct-Dec 20	Jan-March 21	April-June 21	July-Sept 21	Oct-Dec 21
(1)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
p_{11}	0.8195	0.9740	0.9799	0.9701	0.9263	0.9736	0.9746
p_{12}	0.1420	0.0191	0.0124	0.0179	0.0575	0.0155	0.0164
p_{13}	0.0385	0.0069	0.0077	0.0119	0.0162	0.0109	0.0090
p_{21}	0.0794	0.4056	0.2667	0.1846	0.1167	0.3218	0.2121

Period	**April-June 20**	**July-Sept 20**	**Oct-Dec 20**	**Jan-March 21**	**April-June 21**	**July-Sept 21**	**Oct-Dec 21**
p_{22}	0.8571	0.5594	0.6889	0.7692	0.8333	0.6437	0.7273
p_{23}	0.0635	0.0350	0.0444	0.0462	0.0500	0.0345	0.0606
p_{31}	0.0153	0.0676	0.0227	0.0189	0.0191	0.0297	0.0226
p_{32}	0.0153	0.0178	0.0114	0.0113	0.0076	0.0112	0.0113
p_{33}	0.9693	0.9146	0.9659	0.9698	0.9733	0.9591	0.9660

Table 5. Transition Probabilities p_{ij}'s of Female Aged 15 Years and Above in CWS during Selected Quarters for Females in the Urban Sector of India.

Period	**July-Sept 18**	**Oct-Dec 18**	**Jan-March 19**	**April-June 19**	**July-Sept 19**	**Oct-Dec 19**	**Jan-March 20**
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p_{11}	0.8935	0.8817	0.8953	0.8817	0.9123	0.9704	0.9050
p_{12}	0.0118	0.0178	0.0058	0.0178	0.0058	0.0091	0.0200
p_{13}	0.0947	0.1006	0.0988	0.1006	0.0819	0.0205	0.0750
p_{21}	0.0833	0.0455	0.1304	0.0455	0.0909	0.1579	0.0476
p_{22}	0.7083	0.6818	0.6087	0.6818	0.6364	0.7632	0.7619
p_{23}	0.2083	0.2727	0.2609	0.2727	0.2727	0.0789	0.1905
p_{31}	0.0223	0.0173	0.0174	0.0173	0.0235	0.0153	0.0141
p_{32}	0.0074	0.0037	0.0075	0.0037	0.0050	0.0057	0.0039
p_{33}	0.9703	0.9790	0.9752	0.9790	0.9715	0.9790	0.9820

Table 5 (Contd.). Transition Probabilities p_{ij}'s of Females Aged 15 Years and Above in CWS during Selected Quarters of Females in the Urban Areas of India.

Period	**April-June 20**	**July-Sept 20**	**Oct-Dec 20**	**Jan-March 21**	**April-June 21**	**July-Sept 21**	**Oct-Dec 21**
(1)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
p_{11}	0.7413	0.9161	0.9379	0.9227	0.8505	0.9017	0.9106
p_{12}	0.0945	0.0129	0.0056	0.0166	0.0412	0.0116	0.0056
p_{13}	0.1642	0.0710	0.0565	0.0608	0.1082	0.0867	0.0838
p_{21}	0.0435	0.1951	0.1250	0.1111	0.0870	0.1786	0.0870
p_{22}	0.8261	0.6829	0.7500	0.7778	0.7826	0.6786	0.7391
p_{23}	0.1304	0.1220	0.1250	0.1111	0.1304	0.1429	0.1739
p_{31}	0.0052	0.0199	0.0076	0.0101	0.0064	0.0125	0.0113
p_{32}	0.0039	0.0025	0.0025	0.0025	0.0038	0.0025	0.0038
p_{33}	0.9910	0.9776	0.9899	0.9874	0.9898	0.9850	0.9850

These tables separately for males, females and persons provide dimensional idea of retention and loss as well as retention and gain in activity status based on p_{11}, p_{22} and p_{33} i.e. employed, un-employed and out of labour force. It reveals that retention in the employment category was lowest in April-June 2020 and rise in un-employment activity status and out of labour force activity status, which indicates the impact of nationwide lockdown from 24th March 2020 for arresting COVID-19 outbreak, appeared as a national crisis.

Table 6. Transition Probability (p_{22}) of Persons, Males, and Females in the Un-employed Activity Status in a Quarter Compared with the Year Ago Same Quarter, Two Years Ago Same Quarter and Three Years Ago Same Quarter.

Person	**July-Sept 18**	**July-Sept 19**	**July-Sept 20**	**July-Sept 21**	**Oct-Dec 18**	**Oct-Dec 19**	**Oct-Dec 20**	**Oct-Dec 21**
p_{22}	0.6744	0.6829	0.5978	0.6379	0.6977	0.7632	0.7049	0.7333
		Rise	Fall	Fall*		Rise	Stag*	Rise
Male	**July-Sept 18**	**July-Sept 19**	**July-Sept 20**	**July-Sept 21**	**Oct-Dec 18**	**Oct-Dec 19**	**Oct-Dec 20**	**Oct – Dec 21**
p_{22}	0.6508	0.7000	0.5594	0.6437	0.6825	0.8000	0.6889	0.7273
		Rise	Fall	Fall*		Rise	Stag	Rise
Female	**July-Sept 18**	**July-Sept 19**	**July-Sept 20**	**July-Sep 21**	**Oct-Dec 18**	**Oct-Dec 19**	**Oct-Dec 20**	**Oct-Dec 21**
p_{22}	0.7083	0.6364	0.6829	0.6786	0.6818	0.7632	0.75	0.7391
		Fall	Fall	Fall		Rise	Rise	Rise

Table 6 (contd.). Transition Probability (p_{22}) of Persons, Males and Females in the Un-employed Activity Status in a Quarter Compared with the Year Ago Same Quarter, Two Years Ago Same Quarter and Three Years Ago Same Quarter.

Person	**Jan-March 19**	**Jan-March 20**	**Jan-March 21**	**April – June 19**	**April-June 20**	**April-June 21**
p_{22}	0.6667	0.7632	0.5538	0.6977	0.8605	0.8095
		Rise	Rise		Rise	Rise
Male	**Jan-March 19**	**Jan-March 20**	**Jan-March 21**	**April - June 19**	**April-June 20**	**April-June 21**
p_{22}	0.6618	0.7778	0.7692	0.7031	0.8571	0.8333
		Rise	Rise		Rise	Rise
Female	**Jan-March 19**	**Jan-March 20**	**Jan-March 21**	**April - June 19**	**April-June 20**	**April-June 21**
p_{22}	0.6087	0.7619	0.7778	0.6818	0.8261	0.7826
		Rise	Rise		Rise	Rise

Note: * denotes marginal

Further, gender-specific transition probability in selected quarterly periods compared with the corresponding quarter of previous period, two years back and three years back shows that the probability of persons in the same quarter in unemployed activity status in a quarter has either remained stagnant, risen or fallen.

8. Measures of Mobility

The two measures of mobility in activity status (*D* and *J*), stated in this paper i.e. Joshi and Singh's *D* measure (1977) and Joshi's *J* measure (2021) based on transition probabilities given in Tables-3, 4 and 5 for males, females and persons respectively are presented in Table-7.

Table 7. Gender-Specific Measures of Mobility in Selected Quarters in the Urban Sector of India.

Quarterly Survey Period	Males		Females		Persons	
	D Measure	*J* Measure	*D* Measure	*J* Measure	*D* Measure	*J* Measure
July-Sept 2018	0.5272	0.0990	0.5509	0.1298	0.5253	0.0980
Oct -Dec 2018	0.5027	0.0902	0.5845	0.1484	0.5372	0.0993
Jan-March 2019	0.4988	0.0916	0.5954	0.1449	0.5209	0.0955
April-June 2019	0.4823	0.0858	0.5475	0.1352	0.5030	0.0902
July-Sept 2019	0.4944	0.0941	0.5624	0.1327	0.5161	0.0968
Oct-Dec 2019	0.3904	0.0618	0.4220	0.0831	0.4059	0.0658
Jan-March 2020	0.4265	0.0685	0.4816	0.1031	0.4563	0.0756
April-June 2020	0.5073	0.0859	0.5809	0.1583	0.5181	0.0983
July-Sept 2020	0.5245	0.1293	0.5473	0.1160	0.5072	0.1065
Oct-Dec 2020	0.3972	0.0639	0.4583	0.0875	0.4058	0.0645
Jan-March 2021	0.5110	0.0932	0.6590	0.1834	0.5233	0.0943
April-June 2021	0.4153	0.0622	0.5308	0.1202	0.4490	0.0723
July-Sept 2021	0.4556	0.0909	0.5539	0.1221	0.4834	0.0913
Oct-Dec 2021	0.4275	0.0712	0.4897	0.1084	0.4531	0.0779

It provides gender-specific mobility in activity status of persons from current weekly activity status in one quarterly survey period to current weekly activity status in another quarterly survey period. The table reveals gender disparity and inter-temporal disparity in mobility measures in different quarterly periods.
Joshi and Singh's *D* measure of mobility in activity status for males depict decline from 0.5272 in quarter ending September 2018 compared to 0.3904 in quarter ending December 2019. The measure attains maximum rise in the quarter ending March 2021. It came down to the level of 0.4275 in the quarter ending December 2021.

The *J* measure of mobility in activity status for males in Table-7 depicts that mobility in activity status for males has declined from 0.0990 in July–Sept. 2018 to 0.0618 in Oct-Dec 2019 then rose to level of 0.0932 in Jan-March 2021 and fall to the level of 0.0712. Scenario is the same for females except in their dimension. The fall in *D* measure was 0.5509 for females in July-Sept 2018 to 0.4220 in Oct-Dec 2019 and then rise to the level of 0.6590 in Jan-March 2021 in the other quarters given in Table-8 and for measure *J* given in Table-9 compared to the quarter ending September 2018.

Gender-specific *D* measure and *J* measure of mobility in activity status in selected quarterly periods for males, females, and persons compared with the corresponding quarter a year ago, two years ago and three years ago are presented in Table-8 and Table-9 respectively. It shows that for persons, *D* measure of mobility (0.4834) in quarter ending September 2021 compared to *D* measure (0.5072) in quarter ending September 2020, *D* measure of mobility (0.5161) in quarter ending September 2019 and *D* measure of mobility (0.5253) in quarter ending September 2018 has come down. For males, it has come down from the level of 0.5272 in quarter ending September 2018 to the level of 0.4556 in quarter ending September 2021. However, *D* measure for females (0.5539) in quarter ending September 2021 shows marginal increase compared to the level of 0.5509 in quarter ending September 2018. Dimension of Mobility is more in females compared to males in the selected quarterly periods. Following the same conventions in the other quarters given in Table-8 and for measure *J* given in Table-9, there is either rise or fall in dimension of mobility in selected quarters.

Figure 1. Gender-Specific *D* Measure of Mobility in Selected Quarters.

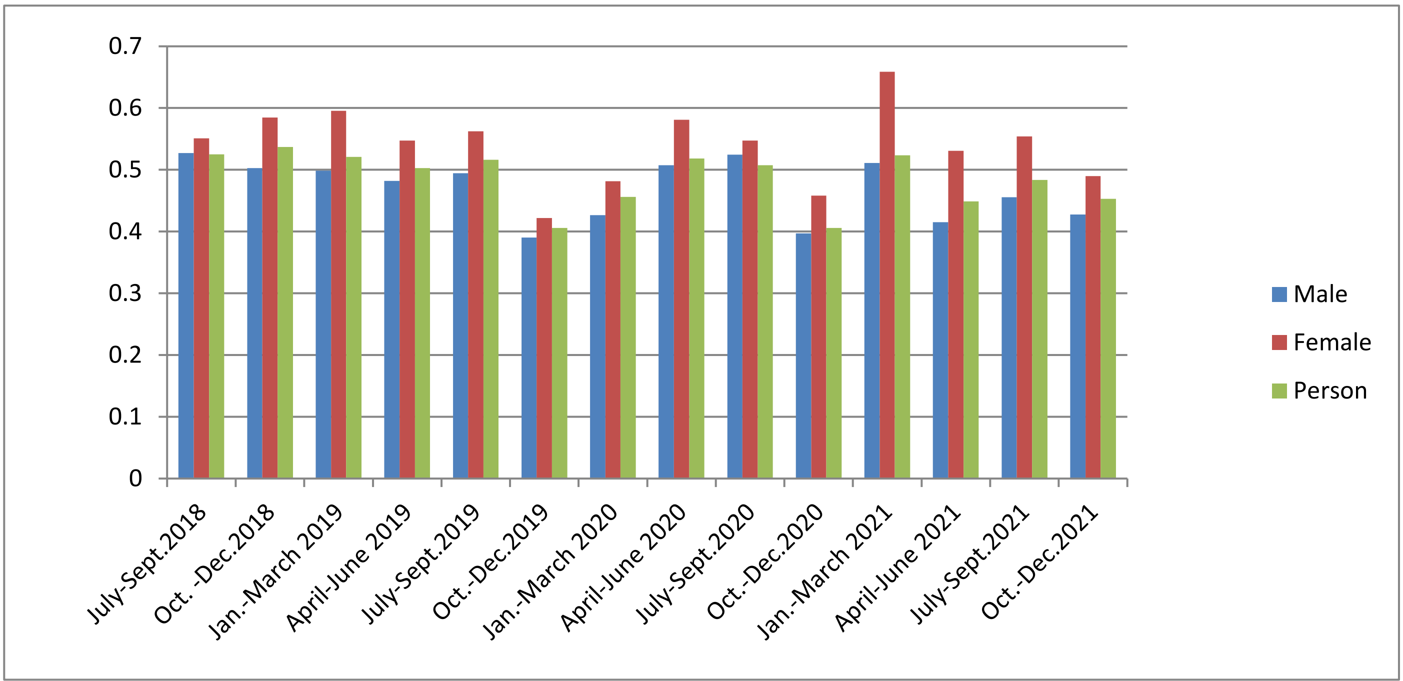


Figure 2. Gender-Specific *J* Measure of Mobility in Selected Quarters.

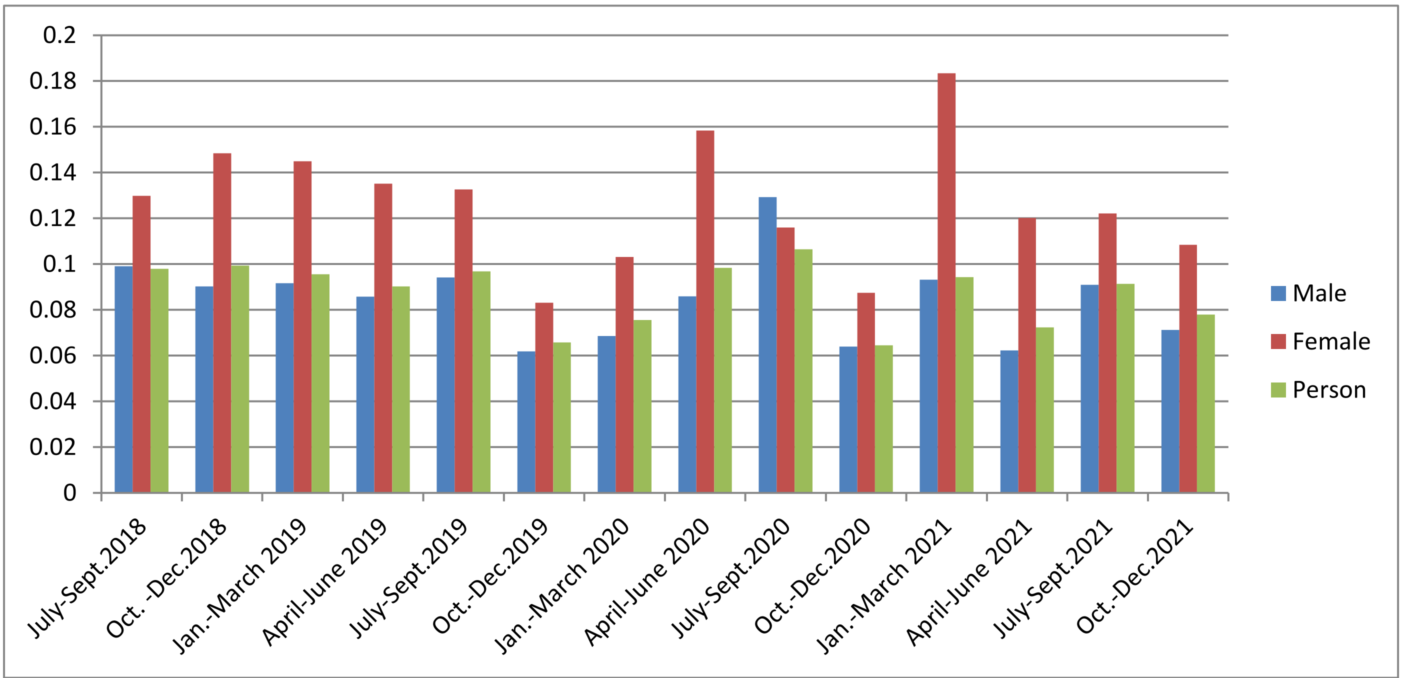


Table 8. *D* Measure of Mobility of Persons, Males, and Females Compared with the Year Ago Same Quarter, Two Year Ago Same Quarter and Three Year Ago Same Quarter.

***D* Measure of Mobility**							
Period	**Person**	**Male**	**Female**	**Period**	**Person**	**Male**	**Female**
July-Sept 2018	0.5253	0.5272	0.5509	Oct-Dec2021	0.4531	0.4275	0.4897
July-Sept 2019	0.5161	0.4944	0.5624	Jan-March 2019	0.5209	0.4988	0.5954
July-Sept 2020	0.5072	0.5245	0.5473	Jan-March 2020	0.4563	0.4265	0.4816
July-Sept 2021	0.4834	0.4556	0.5539	Jan-March 2021	0.5233	0.5110	0.6590
Oct-Dec 2018	0.5372	0.5027	0.5845	April-June 2019	0.5030	0.4823	0.5475
Oct-Dec 2019	0.4059	0.3904	0.4220	April-June 2020	0.5181	0.5073	0.5809
Oct-Dec 2020	0.4058	0.3972	0.4583	April-June 2021	0.4490	0.4153	0.5308

Table 9. *J* Measure of Mobility of Persons, Males, and Females in a Quarter Compared with the Year Ago Same Quarter, Two Year Ago Same Quarter and Three Year Ago Same Quarter.

Period	***J* Measure of Mobility**						
	Person	**Male**	**Female**	**Period**	**Person**	**Male**	**Female**
July-Sept 2018	0.0980	0.0990	0.1298	Oct-Dec 2021	0.0779	0.0712	0.1084
July-Sept 2019	0.0968	0.0941	0.1327	Jan-March 2019	0.0955	0.0916	0.1449
July-Sept 2020	0.1065	0.1293	0.1160	Jan-March 2020	0.0756	0.0685	0.1031
July-Sept 2021	0.0913	0.0909	0.1221	Jan-March 2021	0.0943	0.0932	0.1834
Oct-Dec 2018	0.0993	0.0902	0.1484	April-June 2019	0.0902	0.0858	0.1352
Oct-Dec 2019	0.0658	0.0618	0.0831	April-June 2020	0.0983	0.0859	0.1583
Oct-Dec 2020	0.0645	0.0639	0.0875	April-June 2021	0.0723	0.0622	0.1202

9. Sex Ratio in Measures of Mobility

Sex ratio (F/M) here is the ratio of mobility measure of females to mobility measure of males expressed in 1000. Normalised sex ratio here is the ratio of difference in mobility measure for females and mobility measure for males with mobility measure for males i.e. (F-M)/M. It has been presented in Table 10. However, it shows a rising trend in favour of females. Normalised sex ratio (percentage) based on these two measures also shows a rising trend in favour of females. Sex ratio and normalised sex ratio for both the measures was maximum in the quarter ending March 2021.

Table 10. Sex Ratio (F/M) in Measures of Mobility per Thousand and Normalised Sex Ratio (%) in Mobility Measures based on PLFS at all India Level in the Urban Sector.

Period	**Sex Ratio**		**100*Normalised Sex Ratio**	
	D	***J***	***D***	***J***
July-Sept 2018	1,045	1,311	4.50	31.11
Oct-Dec 2018	1,163	1,645	16.27	64.52
Jan-March 2019	1,194	1,582	19.37	58.19
April-June 2019	1,135	1,576	13.52	57.58
July-Sept 2019	1,138	1,410	13.75	41.02

Period	Sex Ratio		100*Normalised Sex Ratio	
	D	***J***	***D***	***J***
Oct-Dec 2019	1,081	1,345	8.09	34.47
Jan-March 2020	1,129	1,505	12.92	50.51
April-June 2020	1,145	1,843	14.51	84.28
July-Sept 2020	1,043	897	4.35	-10.29
Oct-Dec 2020	1,154	1,369	15.38	36.93
Jan-March 2021	1,290	1,968	28.96	96.78
April-June 2021	1,278	1,932	27.81	93.19
July-Sept 2021	1,216	1,343	21.58	34.32
Oct-Dec 2021	1,145	1,522	14.55	52.25

Sex ratio and normalised sex ratio of mobility measures *D* and *J* in a quarter compared with the same quarter a year ago, two years ago and three years ago has been presented in Table-11 and Table-12. It shows that the highest sex ratio for *D* measure (1,290) was in the quarter ending March 2021. Similar scenario has been observed for *J* measure (1,968) in the quarter ending March 2021.

Table 11. Sex Ratio of Mobility Measure *D* and *J* in a Quarter Compared with the Year Ago Same Quarter, Two Years Ago Same Quarter and Three Years Ago Same Quarter.

Sex Ratio of Mobility Measure *D*				Sex Ratio of Mobility Measure *J*			
Period	**Sex Ratio**	**Period**	**Sex Ratio**	**Period**	**Sex Ratio**	**Period**	**Sex Ratio**
July-Sept 2018	1,045	Oct-Dec 2021	1,145	July-Sept 2018	1,311	Oct-Dec 2021	1,522
July-Sept 2019	1,138	Jan-March 2019	1,194	July-Sept 2019	1,410	Jan-March 2019	1,582
July-Sept 2020	1,043	Jan-March 2020	1,129	July-Sept 2020	897	Jan-March 2020	1,505
July-Sept 2021	1,216	Jan-March 2021	1,290	July-Sept 2021	1,343	Jan-March 2021	1,968
Oct-Dec2018	1,163	April-June 2019	1,135	Oct-Dec 2018	1,645	April-June 2019	1,576
Oct-Dec 2019	1,081	April-June 2020	1,145	Oct-Dec 2019	1,345	April-June 2020	1,843
Oct-Dec 2020	1,154	April-June 2021	1,278	Oct-Dec 2020	1,369	April-June 2021	1,932

Normalised sex ratio of mobility measures *D* and *J* in a quarter to remain in the same quarter in the year ago, two years ago and three years ago presented in Table-12 shows that the highest normalised sex ratio for *D* measure (29) was in the quarter ending March 2021. In the same quarter, *J* measure was also maximum (97).

Table 12. Normalised Sex Ratio of Mobility Measure *D* and *J* in a Compared with the Year Ago Same Quarter, Two Years Ago Same Quarter and Three Years Ago Same Quarter.

Normalised Sex Ratio of Mobility Measure *D*				Normalised Sex Ratio of Mobility Measure *J*			
Period	**NSR**	**Period**	**NSR**	**Period**	**NSR**	**Period**	**NSR**
July-Sept 2018	4	Oct-Dec 2021	15	July-Sept 2018	31	Oct-Dec 2021	52
July-Sept 2019	14	Jan-March 2019	19	July-Sept 2019	41	Jan-March2019	58

Normalised Sex Ratio of Mobility Measure *D*				Normalised Sex Ratio of Mobility Measure *J*			
Period	**NSR**	**Period**	**NSR**	**Period**	**NSR**	**Period**	**NSR**
July-Sept 2020	4	Jan-March 2020	13	July-Sept 2020	-10	Jan-March 2020	51
July-Sept 2021	22	Jan-March 2021	29	July-Sept 2021	34	Jan-March 2021	97
Oct-Dec 2018	16	April-June 2019	14	Oct-Dec 2018	65	April-June 2019	58
Oct-Dec 2019	8	April-June 2020	15	Oct-Dec 2019	34	April-June 2020	84
Oct-Dec 2020	15	April-June 2021	28	Oct-Dec 2020	37	April-June 2021	93

10. Inter-Temporal (Quarterly) Changes in Mobility (Percent)

Quarterly changes in mobility separately for males and females based on previously mentioned measures of mobility have been presented in Table-13. Negative sign in difference shows that mobility in successive quarters is less compared to its previous quarter. Thus, for *D* measure the difference in mobility between October-December 2021 and July-September 2021 is 3.0 percent for persons, 2.8 percent for males and 6.4 percent for females. For *J* measure the difference in mobility between October-December 2021 and July-September 2021 is 3.0 percent for persons, 2.8 percent for males and 6.4 percent for females. For *J* measure the difference in mobility between October-December 2021 and July-September 2021 is 1.3 percent for persons, 2.0 percent for males and 1.4 percent for females.

Table 13. Quarterly Changes in Mobility of Males and Females in the Urban Areas at the All-India Level based on PLFS.

Period *	Male		Female		Persons	
	D	**J**	**D**	**J**	**D**	**J**
Between Q2 and Q1	-0.025	-0.009	0.034	0.019	0.012	0.001
Between Q 3 and Q 2	-0.004	0.001	0.011	-0.004	-0.016	-0.004
Between Q 4 and Q 3	-0.017	-0.006	-0.048	-0.010	-0.018	-0.005
Between Q 5 and Q 4	0.012	0.008	0.015	-0.002	0.013	0.007
Between Q 6 and Q 5	-0.104	-0.032	-0.140	-0.050	-0.110	-0.031
Between Q 7 and Q 6	0.036	0.007	0.060	0.020	0.050	0.010
Between Q 8 and Q 7	0.081	0.017	0.099	0.055	0.062	0.023
Between Q 9 and Q 8	0.017	0.043	-0.034	-0.042	-0.011	0.008
Between Q 10 and Q 9	-0.127	-0.065	-0.089	-0.029	-0.101	-0.042
Between Q 11and Q 10	0.114	0.029	0.201	0.096	0.118	0.030
Between Q 2 and Q 11	-0.096	-0.031	-0.128	-0.063	-0.074	-0.022
Between Q13 and Q 12	0.040	0.029	0.023	0.002	0.034	0.019
Between Q 14 and Q 13	-0.028	-0.020	-0.064	-0.014	-0.030	-0.013

Note: **Survey Period: Q1: July-Sept 2018; Q2: Oct-Dec 2018; Q3: Jan-March 2019; Q4: April-June 2019; Q5: July-Sept 2019; Q6: Oct-Dec 2019; Q7: Jan-March 2020; Q8: April-June 2020; Q9: July- Sept 2020; Q10: Oct-Dec 2020; Q11: Jan-March 2021; Q12: April-June2021; Q13: July-Sept2021; Q14: Oct-Dec 2021.*

Table 14. Normalised Quarterly Change for Mobility Measures in Urban Areas of India.

Period*	Male		Female		Persons	
	D	***J***	***D***	***J***	***D***	***J***
Between Q2 and Q1	-4.65	-8.89	6.10	14.33	2.27	1.33
Between Q3 and Q2	-0.78	1.55	1.86	-2.36	-3.03	-3.83
Between Q4 and Q3	-3.31	-6.33	-8.05	-6.69	-3.44	-5.55
Between Q5 and Q4	2.51	9.67	2.72	-1.85	2.60	7.32
Between Q6 and Q5	-21.04	-34.33	-24.96	-37.38	-21.35	-32.02
Between Q7 and Q6	9.25	10.84	14.12	24.07	12.42	14.89
Between Q8 and Q7	18.94	25.40	20.62	53.54	13.54	30.03
Between Q9 and Q8	3.39	50.52	-5.78	-26.72	-2.10	8.34
Between Q10 and Q9	-24.27	-50.58	-16.26	-24.57	-19.99	-39.44
Between Q11 and Q10	28.65	45.85	43.79	109.60	28.96	46.20
Between Q12 and Q11	-18.73	-33.26	-19.46	-34.48	-14.20	-23.33
Between Q13 and Q12	9.70	46.14	4.35	1.61	7.66	26.28
Between Q14 and Q13	-6.17	-21.67	-11.59	-11.22	-6.27	-14.68

Quarterly changes in the year, two years ago and three years ago for mobility (percent) in activity status of persons, males and females for measures *D* and *J* have been presented in Table-15 and Table-16.

Table 15. Changes in Mobility Measure *D* in a Quarter to Remain in the Same Quarter Compared with the Year Ago Same Quarter, Two Years Ago Same Quarter and Three Years Ago Same Quarter.

Quarterly Changes in *D* Measure							
Period	**Person**	**Male**	**Female**	**Period**	**Person**	**Male**	**Female**
July-Sept 2018	0.5253	0.5272	0.5509	Oct-Dec 2021	0.4531	0.4275	0.4897
July-Sept 2019	0.5161	0.4944	0.5624	Jan-March 2019	0.5209	0.4988	0.5954
July-Sept 2020	0.5072	0.5245	0.5473	Jan-March 2020	0.4563	0.4265	0.4816
July-Sept 2021	0.4834	0.4556	0.5539	Jan-March 2021	0.5233	0.5110	0.6590
Oct-Dec 2018	0.5372	0.5027	0.5845	April-June 2019	0.5030	0.4823	0.5475
Oct-Dec 2019	0.4059	0.3904	0.4220	April-June 2020	0.5181	0.5073	0.5809
Oct-Dec 2020	0.4058	0.3972	0.4583	April-June 2021	0.4490	0.4153	0.5307

Table 16. Changes in Mobility Measure *J* in a Quarter to Remain in the Same Quarter Compared with the Year Ago Same Quarter, Two Years Ago Same Quarter and Three Years Ago Same Quarter.

Quarterly Changes in *J* Measure							
Period	**Person**	**Male**	**Female**	**Period**	**Person**	**Male**	**Female**
July-Sept 2018	0.0980	0.0990	0.1298	Oct-Dec 2021	0.0779	0.0712	0.1084
July-Sept 2019	0.0968	0.0941	0.1327	Jan-March 2019	0.0955	0.0916	0.1449
July-Sept 2020	0.1065	0.1293	0.1160	Jan-March 2020	0.0756	0.0685	0.1031

Quarterly Changes in *J* Measure							
Period	**Person**	**Male**	**Female**	**Period**	**Person**	**Male**	**Female**
July-Sept 2021	0.0913	0.0909	0.1221	Jan-March 2021	0.0943	0.0932	0.1834
Oct-Dec 2018	0.0993	0.0902	0.1484	April-June 2019	0.0902	0.0858	0.1352
Oct-Dec 2019	0.0658	0.0618	0.0831	April-June 2020	0.0983	0.0859	0.1583
Oct-Dec 2020	0.0645	0.0639	0.0875	April-June 2021	0.0723	0.0622	0.1200

It reveals that for *D* measure of mobility, there is fall (0.449) for persons, (0.4153) for males and 0.53079 for females in April-June 2021 compared to (0.5253) for persons, (0.5272) for males and 0.5509 for females in July-Sept 2018. For measure *J*, there is a fall is in April-June 2021 to 0.0723 for persons, 0.0622 for males, 0.1200 for females respectively from 0.098 for persons, 0.099 for males, 0.1298 for females in July-Sept. 2018.

11. Conclusion

Based on results presented in this paper, the following conclusions have been drawn. Mobility in activity status of females towards immobility in the urban sector of India is higher compared to mobility in activity status of males in the selected quarterly periods. Further, a rising trend was seen in regard to mobility in activity status of females in all the selected quarters. For males, the scenario was opposite. Declining trend was seen in all the selected quarters except quarter ending June 2019, which may be attributed to low level of female literacy and women's engagements in household activities. Widening or narrowing inter-temporal gaps over a period of time calls for action-oriented programmes and policies for improving un-employment and under-employment situations.

12. Policy Implications of Results

Dimensional ideas based on measures of employment, under-employment and un- employment are important for the welfare of the people. However, available measures based on rates and ratios for small differences using quinquennial large-scale National Sample Survey data on employment, under- employment and un-employment suffer from sampling variability of the point estimates. Conclusions drawn from the point estimates may therefore not be real and subject to fluctuations of Sampling. Against this, knowledge on success or failures of policy and programmes for reducing under-employment and un-employment in a short period of time from the point of labour force dynamics for taking corrective actions is important. Trend analysis of *J* measure on quarterly basis will, therefore, be helpful to programme implementers and policy formulators for making timely intervention in programmes and policies, if required.

Acknowledgment

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

References

1. Bartholomew, D.J. (1967). Stochastic Models for Social Processes. *John Wiley*.
2. Government of India, MOS & PI, NSSO (2001). Concepts and definitions used in NSS. *Golden Jubilee of NSSO*, Section 4, pp 38-55.
3. Government of India, MOS & PI, NSSO (2019). Periodic Labour Survey (July 2017 – June 2018). *Annual Report, PLFS.*
4. Government of India, MOS & PI, NSO (Nov.2019). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, July- September 2018.
5. Government of India, MOS & PI, NSO (May2019). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, October-December 2018.
6. Government of India, MOS & PI, NSO (2019). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, January- March 2019.
7. Government of India, MOS & PI, NSO (June 2020). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, April-June 2019.
8. Government of India, MOS & PI, NSO (2020). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, July- September 2019.
9. Government of India, MOS & PI, NSO (2020). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, October-December 2019.
10. Government of India, MOS & PI, NSO (2020). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, January- March 2020.
11. Government of India, MOS & PI, NSO (2020). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, April-June 2020.
12. Government of India, MOS & PI, NSO (August 2021). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, July- September 2020.
13. Government of India, MOS & PI, NSO (Sept. 2021). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, October-December 2020.
14. Government of India, MOS & PI, NSO (Nov.2021). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, January- March 2021.
15. Government of India, MOS & PI, NSO (March 2022). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, April-June 2021.
16. Government of India, MOS & PI, NSO (March 2022). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, July- September 2021.
17. Government of India, MOS & PI, NSO (2021). Periodic Labour Force Survey (PLFS). *Quarterly Bulletin*, October- December 2021.
18. Joshi, D.D and B.M. Singh (1977). A measure of intergeneration mobility. *Bull. of Inst. Stat*. XLVII, Book 4.
19. Joshi, P.D (2021). Measure of Dynamic Mobility (J) for Changes in Employment Activity Category based on Transition Probability Matrix. *Sarvekshana*, The Journal of NSSO, Vol.110 and vol.111 (combined).
20. Kundu, Amitabh and P.C. Mohanan (2019). Looking Beyond India's Headline Un-employment number. https://www.livemint.com /politics/policy, !3 August, 2019.

21. Lipset, S.M and R. Bendix (1959). Social Mobility in Industrial Society, Berkley. *University of California Press*.
22. Matras, J. (1960). Comparison of intergenerational mobility patterns, An application of the formal theory of social mobility. *Population Studies*, 14: 2, pp.163-169.
23. Mukherjee, S.P. and R. Basu (1979). Measure of Social and Occupational Mobility. *Demography India*, Vol. XIII, pp. 23.
24. Mukherjee, S.P. and Chattopadhyay, A.K. (1956). Measures of mobility and some associated inferences problems. *Demography India*, 15:269-280.
25. Prais, S.J. (1955). The formal theory of Social mobility. *Population Studies*, Vol. IX, Issue No.1.

Annexure

Important Concepts and Definitions

National Sample Survey on employment un-employment follows clearly defined stabilized concepts given in instructions to field staff Vol-1 as well as in NSS reports on employment un-employment. They are also available in Golden Jubilee publication (2001) of NSSO entitled "Concepts and definitions used in NSS", Section 4, *pp* 38-55. Accordingly, employed are those who work for pay, profit, or family work by gainful activities, i.e. activities that add value to national product. Un-employed are those who are not employed but seeking or available for work. Labour Force means employed and unemployed together. Thus, those who are neither working nor available for work are 'Not in Labour Force'. Further details on the status of activity on which a person spent relatively longer time of the preceding 365 days prior to the date of survey was considered the Principal Usual Activity Status (PUS) of the person. A person who pursued in a subsidiary capacity some gainful activity as well along with their principal usual activity (non-gainful) was considered to be usually working in a Subsidiary Capacity (SUS). Combinations of these two groups constitute all workers in Usual Status (US). The Current Weekly Status (CWS) of labour force rests on longer time of the preceding 7 days prior to the date of survey. The detailed activity statuses under each of the three broad activity statuses (viz., employed', 'unemployed' and 'not in labour force') and the corresponding codes used in the survey are as under:

Code description

Working (or employed)

Self-employed

11 Worked in household enterprises (self-employed) as own-account worker

12 Worked in household enterprises (self-employed) as an employer

21 Worked in household enterprises (self-employed) as helper

Regular wage/ salaried employee

31 Worked as regular wage/salaried employee

Casual labour

41 Worked as casual labour in public works other than MGNREG public works

42 Worked as casual labour in Mahatma Gandhi NREG public works

51 Worked as casual labour in other types of works

61 Did not work owing to sickness though there was work in household enterprise

62 Did not work owing to other reasons though there was work in household enterprise

71 Did not work owing to sickness but had regular salaried/wage employment Seeking work (or unemployed)

72 Did not work owing to other reasons but had regular salaried/wage employment not working but seeking/available for work (or unemployed)

81 Sought work or did not seek but was available for work (for usual status approach)

81 Sought work (for current weekly status approach)

82 Did not seek but was available for work (for current weekly status approach)

Neither working nor available for work (or not in labour force)

91 Attended educational institutions

92 Attended to domestic duties only

93 Attended to domestic duties and was also engaged in free collection of goods (vegetables, roots, firewood, cattle feed, etc.), sewing, tailoring, weaving, etc. for household use

94 Reinters, pensioners, remittance recipients, etc.
95 Not able to work owing to disability
97 Others (including beggars, prostitutes, etc.)
98 Did not work owing to sickness (for casual workers only)
99 Children of age 0-4 years

PART-II

Highlights of Reports Released by National Sample Survey Office (NSSO)

(The 'Highlights' are reproduced from related report prepared by Survey Design and Research Division (SDRD) of NSSO. For details, the reader may refer to the related main report.)

SARVEKSHANA

Highlights of Recent Survey Report(s) Released by National Sample Survey Office (NSSO)

1. Periodic Labour Force Survey (PLFS) (2021-2022)

HIGHLIGHTS

Periodic Labour Force Survey (PLFS) 2021-2022

Survey Period

July 2021 to June 2022

Survey Coverage

Surveyed

12,733 First Stage Units (FSUs)	Rural: 6,988 villages Urban: 5,745 urban blocks
1,01,782 Households	55,895 in rural areas 45,887 in urban areas
4,28,525 Persons	2,49,175 in rural areas 1,79,350 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 2021 - June 2022 emerging from PLFS are highlighted below.

A. Households and Population

Percentage of households with major source of income (household type)

rural households (%)

household type

self-employment	*regular wage/salary earning*	*casual labour*	*others*	*all*
54.0	13.8	25.2	7.1	100.0

urban households (%)

household type

self-employment	*regular wage/salary earning*	*casual labour*	*others*	*all*
33.0	43.2	11.3	12.6	100.0

Literacy rate for persons of age 7 years and above

Literacy Rate for persons of age 7 years and above in India: 79.7%

rural	urban
male: 83.5% female: 68.9%	male: 92.4% female: 84.0%

B. 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: 41.3%

rural	urban
male: 56.9% female: 27.2%	male: 58.3% female: 18.8%

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: 42.0%

rural:	urban:
42.6%	40.6%

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: 55.2%

rural:	urban:
57.5%	49.7%

C. Workforce

WPR
in usual status
for persons of
all ages

Worker Population Ratio (WPR) in usual status (ps+ss) in India: 39.6%

rural	urban
male: 54.7%	male: 55.0%
female: 26.6%	female: 17.3%

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: 36.8%

rural:	urban:
38.0%	33.6%

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: 52.9%

rural:	urban:
55.6%	46.6%

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

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

rural male: 58.6	rural female: 67.8	urban male: 39.5	urban female: 39.4

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

rural male: 14.7	rural female: 8.1	urban male: 46.2	urban female: 50.3

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

rural male: 26.8	rural female: 24.1	urban male: 14.3	urban female: 10.3

Industry of work (NIC - 2008) of the workers in usual status (ps+ss)

Some industry of work with share (%) of workers in usual status (ps+ss) in rural areas

Agriculture Sector	rural male: 51.0	rural female: 75.9	rural person: 59.0
Construction Sector	rural male: 16.6	rural female: 5.3	rural person: 13.0
Trade, hotel and restaurant Sector	rural male: 10.6	rural female: 3.7	rural person: 8.4
Manufacturing Sector	rural male: 7.9	rural female: 7.9	rural person: 7.9

Industry of work (NIC - 2008) of the workers in usual status (ps+ss)

Some industry of work with share (%) of workers in usual status (ps+ss) in urban areas

Trade, hotel and restaurant Sector	urban male: 25.2	urban female: 14.8	urban person: 22.8
Manufacturing Sector	urban male: 21.5	urban female: 24.3	urban person: 22.2
Construction Sector	urban male: 12.9	urban female: 3.9	urban person: 10.8
Transport, storage & communications	urban male: 12.5	urban female: 4.6	urban person: 10.7

Occupation (Division of NCO-2015) of workers in usual status (ps+ss)

Some Occupation Divisions (NCO-2015) with share (%) of workers in usual status (ps+ss) in rural areas

Division 6: Skilled Agricultural, Forestry and Fishery Workers	rural male: 40.6	rural female: 57.5
Division 7: Craft and Related Trades Workers	rural male: 8.8	rural female: 5.4
Division 5: Service and Sales Workers	rural male: 8.2	rural female: 4.2

Occupation (Division of NCO-2015) of workers in usual status (ps+ss)

Some Occupation Divisions (NCO-2015) with share (%) of workers in usual status (ps+ss) in urban areas

Division 5: Service and Sales Workers	urban male: 18.6	urban female: 16.7
Division 7: Craft and Related Trades Workers	urban male: 16.1	urban female: 13.6
Division 1: Managers	urban male: 16.3	urban female: 10.0
Division 2: Professionals	urban male: 9.7	urban female: 18.2

Informal Sector

Informal Sector

Percentage of workers in usual status (ps+ss) engaged in informal non-agriculture sector in India:

male: 75.2 | female: 58.4 | person: 71.8

Conditions of employment

Conditions of Employment

Percentage of regular wage/salaried employees in the non-agriculture sector who had no job contract in India

male: 62.9 | female: 59.1 | person: 62.0

Percentage of regular wage/salaried employees in the non-agriculture sector who were not eligible for paid leave in India

male: 50.5 | female: 44.6 | person: 49.2

Percentage of regular wage/salaried employees in the non-agriculture sector who were not eligible for specified social security# in India

male: 52.2 | female: 55.7 | person: 53.0

#: In PLFS, coverage of social security for *regular wage/salaried employees* means whether they are covered under any of the following specified social security benefits or a combination of these benefits which are arranged or for which contribution is made by the employer.

- PF/ pension
- gratuity
- health care / maternity benefits

D. Earnings from employment, hours worked and hours available for additional work

Earnings from employment, hours worked and hours available for additional work

Estimates derived based on

- data collected in first visit schedule in rural areas; and
- data collected in first visit and revisit schedule in urban areas during for each of the survey periods July – September 2021, October- December 2021, January – March 2022 and April – June 2022

Information on earnings collected for

- self-employed persons in current weekly status (CWS) for last 30 days
- regular wage/salaried persons in current weekly status (CWS) for last calendar month
- casual labour during each day of reference week

Range of earnings from employment of regular wage/ salaried employees In CWS

Range of earnings for regular wage/salaried employees in CWS during preceding calendar month in the quarters July – September 2021, October-December 2021, January – March 2022 and April – June 2022

rural	male	₹16.0 thousand - ₹ 16.5 thousand
	female	₹9.8 thousand - ₹ 12.6 thousand
urban	male	₹21.5 thousand - ₹ 22.8 thousand
	female	₹ 17.0 thousand - ₹ 18.0 thousand

Range of earnings from employment by casual labour engaged in work other than public

Average wage earnings per day by casual labour engaged in work other than public works during the reference week of the quarters July – September 2021, October- December 2021, January – March 2022 and April – June 2022

rural	male	₹ 381 - ₹ 393
	female	₹ 258 - ₹ 265
urban	male	₹ 450 - ₹ 483
	female	₹ 317 - ₹ 333

Range of earnings from employment of self-employed workers

Average gross earnings during last 30 days from self-employment work by self-employed workers in CWS in the quarters July – September 2021, October- December 2021, January – March 2022 and April – June 2022

rural	male	₹ 10.7 thousand - ₹ 12.1 thousand
	female	₹ 4.7 thousand - ₹ 4.9 thousand
urban	male	₹ 17.6 thousand - ₹ 19.6 thousand
	female	₹7.4 thousand - ₹ 7.9 thousand

Hours actually worked during the reference week by workers in CWS

Average hours actually worked in a week by a worker in CWS during Jul 2021 – June 2022: 43.7 hours – 44.8 hours

rural:	urban:
41.9 hours – 43.1 hours	48.6 hours – 49.3 hours

Hours available for additiona l work by the workers in CWS

Percentage of workers (range) in CWS who reported that they were available for additional work during July 2021 – June 2022

rural:	urban:
2.4 % -3.8%	1.2 % -1.4%

Hours available for additional work (range) in a week for workers in CWS who reported that they were available for additional work during July 2021 – June 2022

rural:	urban:
10.9 hours -12.2 hours	9.7 hours -11.7 hours

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

Jn-employmen Rate (UR) in usual status for persons of all ages

Unemployment Rate in usual status (ps+ss) for persons of all ages in India: 4.1%

rural	urban
male: 3.8% female: 2.1%	male: 5.8% female: 7.9%

Un-employment 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: 8.6%

rural	urban
8.0%	9.5%

Un-employment 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: 12.4%

rural	urban
male: 11.4% female: 8.5%	male: 15.8% female: 21.6%

F. 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) and PLFS (2021-22)

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 (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

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

Table 2: WPR (in per cent) in usual status (ps+ss) estimated from PLFS (2017-18), PLFS(2018-19), PLFS (2019-20), PLFS (2020-21) and PLFS (2021-22) 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 (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

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

Table 3: Unemployment Rate (in per cent) in usual status (ps+ss) estimated from PLFS (2017-18), PLFS (2018-19), PLFS (2019-20), PLFS (2020-21) and PLFS (2021-22)

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 (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

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

खण्ड-III हिंदी

सर्वेक्षण

राष्ट्रीय प्रतिदर्श सर्वेक्षण कार्यालय की पत्रिका

भाग सं० - PDOS-XXXVIII सं॰ 3 और 4

अंक संख्या 114वां

मार्च, 2023

सत्यमेव जयते

राष्ट्रीय प्रतिदर्श सर्वेक्षण कार्यालय

सांख्यिकी और कार्यक्रम कार्यान्वयन मंत्रालय

भारत सरकार

नई दिल्ली

सर्वेक्षण

भाग सं० - PDOS-XXXVIII सं॰ 3 और 4

एनएसएसओ द्वारा जारी की गई रिपोर्ट की मुख्य बातें

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

मुख्य बातें

आवधिक श्रमबल सर्वेक्षण (पीएलएफएस) 2021-2022

सर्वेक्षण

जुलाई 2021 से जून 2022

सर्वेक्षण कवरेज

सर्वेक्षण किया गया

12,733 फर्स्ट स्टेज यूनिट्स (एफएसयु)	ग्रामीण: 6,988 गांवों नगरीय: 5,745 नगरीय खंडो
1,01,782 परिवारों	55,895 ग्रामीण क्षेत्रों में 45,887 नगरीय क्षेत्रों में
4,28,525 वयक्तियों	2,49,175 ग्रामीण क्षेत्रों में 1,79,350 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 2021 - जून 2022 की अवधि के लिए पीएलएफएस से प्राप्त कुछ मुख्य परिणाम निम्नलिखित हैं।

(क) परिवार एवं जनसंख्या

परिवारों के आय का प्रमुख स्त्रोत का प्रतिशत (परिवारों के प्रकार)

ग्रामीण परिवारों (%)

परिवारों के प्रकार

स्व.नियोजन	नियमित मजदूर/वेतन	आकस्मिक मजदूरी	अन्य	सब
54.0	13.8	25.2	7.1	100.0

नगरीय परिवारों (%)

परिवारों के प्रकार

स्व.नियोजन	नियमित मजदूर/वेतन	आकस्मिक मजदूरी	अन्य	सब
33.0	43.2	11.3	12.6	100.0

साक्षरता दर 7 वर्ष और उससे अधिक उम्र के व्यक्तियों में

भारत में साक्षरता दर 7 वर्ष और उससे अधिक उम्र के व्यक्तियों में: 79.7%

ग्रामीण	नगरीय
पुरूषों में: 83.5%	पुरूषों में: 92.4%
महिलाओं में: 68.9%	महिलाओं में: 84.0%

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

एलएफपीआर सभी उम्र के व्यक्तियों में

भारत में लेबर फोर्स पार्टिसपैशन रेट (एलएफपीआर) सामान्य स्थिति (पीएस+एसएस) में: 41.3%

ग्रामीण	नगरीय
पुरूषों में: 56.9%	पुरूषों में: 58.3%
महिलाओं में: 27.2%	महिलाओं में: 18.8%

एलएफपीआर 15-29 वर्ष उम्र के व्यक्तियों में

भारत के 15-29 वर्ष उम्र के व्यक्तियों में सामान्य स्थिति (पीएस+एसएस) में एलएफपीआर: 42.0%

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
42.6%	40.6%

एलएफपीआर 15 वर्ष एवं उससे अधिक उम्र के व्यक्तियों में

भारत के 15 वर्ष एवं उससे अधिक उम्र के व्यक्तियों में सामान्य स्थिति (पीएस+एसएस) में एलएफपीआर: 55.2%

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
57.5%	49.7%

(ग) कार्यबल

डब्ल्यूपीआर सभी उम्र के व्यक्तियों में

कामगार जनसंख्या अनुपात (डब्ल्यूपीआर) सामान्य स्थिति में (पीएस+एसएस) में: 39.6%

<u>ग्रामीण</u>	<u>नगरीय</u>
पुरूषों में: 54.7%	पुरूषों में: 55.0%
महिलाओं में: 26.6%	महिलाओं में: 17.3%

डब्ल्यूपीआर 15-29 वर्ष उम्र के व्यक्तियों में

15-29 वर्ष उम्र के व्यक्तियों में भारत में डब्ल्यूपीआर सामान्य स्थिति (पीएस+एसएस) में: 36.8%

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
38.0%	33.6%

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: 52.9%

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
55.6%	46.6%

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

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

<u>ग्रामीण</u>	<u>ग्रामीण</u>	<u>नगरीय</u>	<u>नगरीय</u>
पुरूषों में: 58.6	महिलाओं में: 67.8	पुरूषों में: 39.5	महिलाओं में: 39.4

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

<u>ग्रामीण</u>	<u>ग्रामीण</u>	<u>नगरीय</u>	<u>नगरीय</u>
पुरूषों में: 14.7	महिलाओं में: 8.1	पुरूषों में: 46.2	महिलाओं में: 50.3

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

<u>ग्रामीण</u>	<u>ग्रामीण</u>	<u>नगरीय</u>	<u>नगरीय</u>
पुरूषों में: 26.8	महिलाओं में: 24.1	पुरूषों में: 14.3	महिलाओं में: 10.3

सामान्य स्थिति (पीएस+एसएस) में कामगारों का कार्य उद्योग (एनआईसी - 2008)

सामान्य स्थिति (पीएस+एसएस) में कुछ कार्य उद्योग (एनआईसी -2008) में कामगारों का शेयर (%) <u>ग्रामीण क्षेत्रों में</u>

कृषि क्षेत्र में	<u>ग्रामीण</u> पुरूषों में: 51.0	<u>ग्रामीण</u> महिलाओं में: 75.9	<u>ग्रामीण</u> व्यक्तियों में: 59.0
'निर्माण' सेक्टर में	<u>ग्रामीण</u> पुरूषों में: 16.6	<u>ग्रामीण</u> महिलाओं में: 5.3	<u>ग्रामीण</u> व्यक्तियों में: 13.0
ट्रेड, होटल और रेस्टुरेन्ट सेक्टर में	<u>ग्रामीण</u> पुरूषों में: 10.6	<u>ग्रामीण</u> महिलाओं में: 3.7	<u>ग्रामीण</u> व्यक्तियों में: 8.4
'विनिर्माण' क्षेत्र में	<u>ग्रामीण</u> पुरूषों में: 7.9	<u>ग्रामीण</u> महिलाओं में: 7.9	<u>ग्रामीण</u> व्यक्तियों में: 7.9

सामान्य स्थिति (पीएस+एसएस) में कामगारों का कार्य उद्योग (एनआईसी- 2008)

सामान्य स्थिति (पीएस+एसएस) में कुछ कार्य उद्योग (एनआईसी -2008) में कामगारों का शेयर (%) नगरीय क्षेत्रों में

ट्रेड, होटल और रेस्टुरेन्ट सेक्टर में	<u>नगरीय</u> पुरूषों में: 25.2	<u>नगरीय</u> महिलाओं में: 14.8	<u>नगरीय</u> व्यक्तियों में: 22.8
'विनिर्माण' क्षेत्र में	<u>नगरीय</u> पुरूषों में: 21.5	<u>नगरीय</u> महिलाओं में: 24.3	<u>नगरीय</u> व्यक्तियों में: 22.2
'निर्माण' सेक्टर में	<u>नगरीय</u> पुरूषों में: 12.9	<u>नगरीय</u> महिलाओं में: 3.9	<u>नगरीय</u> व्यक्तियों में: 10.8
ट्रैन्स्पर्टैशन, स्टॉरिज एण्ड कम्यूनिकेशन	<u>नगरीय</u> पुरूषों में: 12.5	<u>नगरीय</u> महिलाओं में: 4.6	<u>नगरीय</u> व्यक्तियों में: 10.7

सामान्य स्थिति (पीएस+एसएस) में कामगारों का उपजीविका (एनसीओ 2015 के प्रभाग)

सामान्य स्थिति (पीएस+एसएस) में कुछ उपजीविका (एनसीओ 2015 के प्रभाग) में कामगारों का शेयर (%)

ग्रामीण क्षेत्रों में

प्रभाग 6: कुशल कृषि, वानिकी एवं मत्स्य पालन में युक्त कामगार	ग्रामीण पुरूषों में: 40.6	ग्रामीण महिलाओं में: 57.5
प्रभाग 7 : कारीगरी एवं संबंधित ट्रैड में युक्त कामगार	ग्रामीण पुरूषों में: 8.8	ग्रामीण महिलाओं में: 5.4
प्रभाग 5: सेवाएं एवं विक्रय में युक्त कामगार	ग्रामीण पुरूषों में: 8.2	ग्रामीण महिलाओं में: 4.2

सामान्य स्थिति (पीएस+ एसएस)में कामगारों का उपजीविका (एनसीओ 2015 के प्रभाग)

सामान्य स्थिति (पीएस+एसएस) में कुछ उपजीविका (एनसीओ 2015 के प्रभाग) में कामगारों का शेयर (%)

नगरीय क्षेत्रों में

प्रभाग 5: सेवाएं एवं विक्रय में युक्त कामगार	नगरीय पुरूषों में: 18.6	नगरीय महिलाओं में: 16.7
प्रभाग 7: कारीगरी एवं संबंधित ट्रैड में युक्त कामगार	नगरीय पुरूषों में: 16.1	नगरीय महिलाओं में: 13.6
प्रभाग 1: प्रबंधकों	नगरीय पुरूषों में: 16.3	नगरीय महिलाओं में: 10.0
प्रभाग 2: पेशेवरों	नगरीय पुरूषों में: 9.7	नगरीय महिलाओं में: 18.2

अनौपचारिक क्षेत्र

गैर-कृषि अनौपचारिक क्षेत्र

भारत में गैर-कृषि अनौपचारिक क्षेत्र में नियुक्त सामान्य स्थिति (पीएस+एसएस) में कामगारों का प्रतिशत (%):

पुरुष कामगारों में:	महिला कामगारों में:	व्यक्ति कामगारों में:
75.2	58.4	71.8

सामान्य स्थिति (पीएस+एसएस) में कामगारों के रोजगार की अवस्था

सामान्य स्थिति (पीएस+एसएस) में कामगारों के रोजगार की अवस्था

भारत में, गैर-कृषि क्षेत्र में नियमित मजदूर/वेतन भोगी कर्मचारियों जिनके पास कोई लिखित नौकरी संविदा नहीं था उनका प्रतिशत (%):

पुरूषों में:	महिलाओं में:	व्यक्तियों में:
62.9	59.1	62.0

भारत में, गैर-कृषि क्षेत्र में नियमित मजदूर /वेतन भोगी कर्मचारियों जो वेतन युक्त अबकाश के योग्य नहीं थे उनका प्रतिशत (%):

पुरूषों में:	महिलाओं में:	व्यक्तियों में:
50.5	44.6	49.2

भारत में, गैर-कृषि क्षेत्र में नियमित मजदूर/वेतन भोगी कर्मचारियों जो किसी सामाजिक सुरक्षा# हितलाभ के पात्र नहीं थे उनका प्रतिशत (%):

पुरूषों में:	महिलाओं में:	व्यक्तियों में:
52.2	55.7	53.0

#: पीएलएफएस में नियमित मजदूर /वेतनभोगी कामगारों के लिए सामाजिक सुरक्षा का कवरेज का तात्पर्य यह हैं की क्या वे निम्नलिखित निर्दिष्ट सामाजिक सुरक्षा लाभों या उनके किसी संयोजन के तहत शामिल थे जिसकी व्यवस्था नियोक्ता द्वारा की जाती है या जिसके लिए अंशदान नियोक्ता द्वारा दिया जाता है |

- पीएफ / पेंशन
- ग्रेच्युटी
- हेल्थ केयर / मटर्निटी बेनीफिट्स

(घ) कामगारों के आय, कितने घंटे काम किया एवं अतिरिक्त कार्यों के लिए उपलब्ध घंटे

कामगारों के आय, कितने घंटे काम किया एवं अतिरिक्त कार्यों के लिए उपलब्ध घंटे

एस्टीमटेस आधारित है

- ग्रामीण क्षेत्रों में किए गए अनुसूची के पहले दौरे पर इकट्ठे किए गए आंकड़ों; और
- नगरीय क्षेत्रों में अनुसूची के पहले दौरे पर और पुनः दौरे पर इकट्ठे किए गए आंकड़ों जो जुलाई - सितम्बर 2021, अक्टूबर - दिसंबर 2021, जनवरी - मार्च 2022 एवं अप्रैल - जून 2022 अवधियों के लिए थे

रोजगार से आय पर सुचना इकट्ठी की गयी

- वर्तमान साप्ताहिक स्थिति (सीडब्ल्यूएस) में स्व-रोजगार व्यक्तियों के लिए आय पर सुचना पिछले 30 दिनों के लिए
- वर्तमान साप्ताहिक स्थिति (सीडब्ल्यूएस) में नियमित मजदूर/वतनभोगी व्यक्तियों के लिए आय पर सुचना पूर्ववर्ती केलेण्डर माह के लिए
- आकस्मिक श्रमिक के लिए आय पर सुचना संदर्भ हफ्ते के प्रतिदिन के लिए

वर्तमान साप्ताहिक स्थिति (सीडब्ल्यूएस) में नियमित मजदूर/वेतन भोगी कर्मचारियों के

वर्तमान साप्ताहिक स्थिति (सीडब्ल्यूएस) में नियमित मजदूर/वेतनभोगी कर्मचारियों के बीच, पूर्ववर्ती केलेण्डर माह के दौरान हुई आय की रेंज सर्वेक्षण अवधि के जुलाई - सितम्बर 2021, अक्टूबर - दिसंबर 2021, जनवरी - मार्च 2022 एवं अप्रैल - जून 2022 के बीच में

ग्रामीण	
पुरूषों में	₹ 16.0 हजार - ₹ 16.5 हजार
महिलाओं में	₹ 9.8 हजार - ₹ 12.6 हजार
नगरीय	
पुरूषों में	₹ 21.5 हजार - ₹ 22.8 हजार
महिलाओं में	₹ 17.0 हजार - ₹ 18.0 हजार

आकस्मिक श्रमिक (पब्लिक वर्क्स के अलावा अन्य कार्य में) की औसतन प्रतिदिन आय

आकस्मिक श्रमिक (पब्लिक वर्क्स के अलावा अन्य कार्य में) की औसतन प्रतिदिन की आय सर्वेक्षण अवधि के जुलाई - सितम्बर 2021, अक्टूबर - दिसंबर 2021, जनवरी - मार्च 2022 एवं अप्रैल - जून 2022 के बीच में

ग्रामीण	
पुरूषों में	₹381 - ₹ 393
महिलाओं में	₹ 258 - ₹ 265
नगरीय	
पुरूषों में	₹450 - ₹ 483
महिलाओं में	₹317 - ₹ 333

सीडब्ल्यूएस में स्व-कार्यरत कामगारों दवारा किए गए स्व-कार्यरत कार्य से औसतन कुल आय की रेंज

सीडब्ल्यूएस में स्व-कार्यरत कामगारों द्वारा किए गए स्व-कार्यरत कार्य से औसतन कुल आय की रेंज सर्वेक्षण अवधि के जुलाई - सितम्बर 2021, अक्टूबर - दिसंबर 2021, जनवरी - मार्च 2022 एवं अप्रैल - जून 2022 के बीच में

ग्रामीण	
पुरूषों में	₹ 10.7 हजार - ₹ 12.1 हजार
महिलाओं में	₹ 4.7 हजार - ₹ 4.9 हजार
नगरीय	
पुरूषों में	₹ 17.6 हजार - ₹ 19.6 हजार
महिलाओं में	₹ 7.4 हजार - ₹ 7.9 हजार

सीडब्ल्यूएस में कामगार द्वारा औसतन साप्ताहिक कितने घंटे कार्य किया गया

वर्तमान साप्ताहिक स्थिति (सीडब्ल्यूएस) में सर्वेक्षण अवधि जुलाई 2021 से जून 2022 के दौरान कामगार द्वारा औसतन साप्ताहिक काम किया गया: 43.7 घंटे – 44.8 घंटे

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
41.9 घंटे - 43.1 घंटे	48.6 घंटे - 49.3 घंटे

सीडब्ल्यूएस में कामगारों का अतिरिक्त कार्यों के लिए उपलब्ध समय

वर्तमान साप्ताहिक स्थिति (सीडब्ल्यूएस) में अतिरिक्त कार्य की उपलब्धता दर्ज करवाने वाले कामगारों की प्रतिशत की रेंज सर्वेक्षण अवधि जुलाई 2021 से जून 2022 के दौरान

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
2.4 % -3.8%	1.2 % - 1.4%

सीडब्ल्यूएस में जिन कामगारों ने अतिरिक्त कार्य की उपलब्धता दर्ज करवायी थी उस में एक हफ्ते में अतिरिक्त कार्य की लिए उपलब्ध समय की रेंज सर्वेक्षण अवधि जुलाई 2021 से जून 2022 के दौरान

ग्रामीण क्षेत्रों में:	नगरीय क्षेत्रों में:
10.9 घंटे -12.2 घंटे	9.7 घंटे -11.7 घंटे

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

बेरोजगार दर सभी उम्र के व्यक्तियों पर

भारत में सामान्य स्थिति (पीएस+एसएस) में बरोजगार दर: 4.1%

ग्रामीण	नगरीय
पुरूषों में: 3.8%	पुरूषों में: 5.8%
महिलाओं में: 2.1%	महिलाओं में: 7.9%

बरोजगार दर 15 बर्षो और उससे उपर के उम्र के शिक्षित व्यक्तियों में

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: 8.6%

ग्रामीण क्षेत्रों में	नगरीय क्षेत्रों में
8.0%	9.5%

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

भारत में बरोजगार दर युवा (15-29 वर्ष उम्र के) व्यक्तियों में: 12.4%

ग्रामीण	नगरीय
पुरूषों में: 11.4%	पुरूषों में: 15.8%
महिलाओं में: 8.5%	महिलाओं में: 21.6%

(च). प्रमुख श्रम बल संकेतक की टाइम-सीरीज सामान्य स्थिति (पीएस+एसएस) में पीएलएफएस सें 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) एवं पीएलएफएस (2021-22) से प्राक्कलित

अल-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 (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

2021-22 जुलाई 2021-जून 2022 की अवधि को संदर्भित करता है और इसी तरह 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)** एवं पीएलएफएस **(2021-22)** से प्राक्कलित

अल-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 (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

2021-22 जुलाई 2021-जून 2022 की अवधि को संदर्भित करता है और इसी तरह 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) एवं पीएलएफएस (2021-22) से प्राक्कलित

अल-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 (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

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

BROAD CONTENTS OF THE LAST FOUR ISSUES OF 'SARVEKSHANA'

PDOS XXXVIII No. 3 & 4
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, Survey Coordination Division, National Sample Survey 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