Indian Journal of Agricultural Economics 80: 1 (2025): 201-214
DOI:10.63040/25827510.2025.01.013

RESEARCH NOTE

Evaluating Agricultural Labour Data Challenges: Situation Assessment Survey Data vs Cost of Cultivation of Principal Crops Data

Prachi Bansal[1]*

ABSTRACT

Most national surveys, such as the NSS Employment-Unemployment Surveys and the Periodic Labour Force Surveys, discuss data on workers in agriculture. However, given the limited supply of land and a surplus of labour looking for opportunities to find work, it is critical to study labour use and labour costs in agriculture. Labour absorption or labour use is measured in terms of labour time. There are limited studies on labour use in agriculture, and the two major sources of large-scale national data on this issue are the Situation Assessment Survey (SAS) conducted by the National Statistical Office and the data collected under the Comprehensive Scheme for Cost of Cultivation of Principal Crops (CCPC) by the Ministry of Agriculture. These two datasets provide valuable insights into labour use and labour expenditure statistics, though each has certain limitations. Taking a different approach, this paper focuses on examining the challenges and methodological issues in using these datasets to understand labour use and expenditure patterns in farming rather than presenting these estimates of labour use. Two primary issues are identified using SAS data: the absence of data on the number of labour days used in agriculture and the lack of detailed data on crop-wise expenditure on labour. However, the SAS data does give us an aggregate measure of labour expenditure (or costs) in crop cultivation. This aggregate measure is missing from the CCPC surveys, which collect detailed information on cost and labour days in agriculture, but the data are released crop-wise. Consequently, we have a situation where the SAS data provides an aggregate picture of expenditure on labour in crop cultivation, whereas the CCPC data gives us this information crop-wise. This paper shows that the expenditure on labour so estimated using the SAS grossly underestimates labour costs in agriculture compared to the CCPC data and, thus, underestimates the total costs incurred in crop cultivation. This paper highlights the necessity of improving data collection methods to obtain a more accurate and detailed understanding of labour use and expenditure in agriculture, which is critical for formulating policies for the agricultural sector.

Keywords: Labour use, cost of production data, situation assessment survey data, labour expenditure

JEL codes: C83, D24, Q13, O13

I

INTRODUCTION

The studies in the Economics of Farm Management in India, initiated by the Directorate of Economics and Statistics, Ministry of Agriculture, were the first nation-wide studies that collected data on farm enterprises. These surveys aimed to collect data on the costs of cultivation and returns from various crops to derive reliable estimates of income from agriculture (Surjit, 2017). These studies were started in 1954–55 in six regions and were expanded to other parts of the country in subsequent years. These covered only one or two districts of major states and provided estimates for selected crops. These studies were not continuously carried out in the selected

[1]Jindal School of Government and Public Policy, O.P. Jindal Global University, Sonipat, Haryana – 131 001.
*The author grateful to Prof. Vikas Rawal for his invaluable guidance and support during her doctoral research on labour absorption in Indian agriculture at Jawaharlal Nehru University. Although he has not directly reviewed this paper, his mentorship has laid the foundation of this work.

regions; thus, these do not provide time series data. The FMS data were collected until the year 1972–73.

Although the FMS data had many limitations and were not uniform, scholars extensively used these to discuss the effects of the Green Revolution, adoption of modern varieties, efficiency in production, impact of mechanisation, costs of production, and many other important aspects of agriculture. The debate on farm size and farm productivity in Indian agriculture was primarily based on data from these surveys. Major strands of literature emerged between the 1950s and 1970s based on these surveys. Many studies on labour absorption, expenditure on labour in crop cultivation, efficiency or optimal use of labour input and studies that examined the production conditions in Indian agriculture were carried out extensively (Bardhan, 1973; Bharadwaj, 1974; Gangwar, 1970; Paranjpe, 1958; Patnaik, 1971; Saini, 1969; Sen, 1962; Vaidyanathan, 1978). Numerous significant studies in this area were published during that time in the Indian Journal of Agricultural Economics (Acharya, 1992; Bhalla, 1987; Dantwala, 1987; Gangwar, 1970; Jena, 1957; Prakash, 1962; Saini, 1969).

The studies based on FMS data had limitations in their use for any price policy recommendation nationally or even for a state because of its limited coverage and non-uniformity (Sen and Bhatia, 2004). Thus, it was decided to do away with the Farm Management Studies and introduce a continuous series of data on farm economics. The Comprehensive Scheme for Cost of Cultivation of Principal Crops (CCPC) was planned in 1970–71. Under CCPC, data are collected for major crops in the major States of India. It is used for annual recommendations on agricultural price policy made by the Commission for Agricultural Costs and Prices (CACP). A continuous data series was available for agricultural economists as these studies began. However, due to the focus on price policy recommendations and relative difficulty in using these data, these remained underutilized, particularly to understand the pattern of absorption of labour in agriculture.

Another large-scale data source that records the pattern of labour used in agriculture, viz. hired or family, is the Situation Assessment Survey (SAS) data from the National Statistical Office. The Situation Assessment Survey (SAS) was first conducted in 2003 by the NSSO as part of its 59th round. It was undertaken again in 2013 as part of the 70th round and in 2019 as part of the 77th round. The SAS schedule was designed to collect information on various aspects of farming and other socio-economic characteristics of agricultural households.

For clarity, it is essential to note that this paper does not present estimates of labour absorption based on the two datasets. Instead, it examines the challenges and methodological issues in using these datasets to understand labour expenditure and use in Indian farming. Section 2 describes the Cost of Cultivation of Principal Crops (CCPC) data. Section 3 explains the challenges and methodological issues associated

with using the farm-level CCPC data. Section 4 describes the Situation Assessment Survey data. Section 5 shows that the expenditure on labour estimated using SAS data grossly underestimates labour costs in agriculture compared to the CCPC data, consequently underestimating the total costs incurred in crop cultivation. Section 6 highlights the necessity of improving data collection methods to obtain a more accurate and detailed understanding of labour use and expenditure in agriculture.

II

DATA ON THE COST OF CULTIVATION OF PRINCIPAL CROPS

The Comprehensive Scheme for Cost of Cultivation of Principal Crops (CCPC) was planned in 1970–71. Under CCPC, data on costs and input use are collected for major crops in the major States of India. It is used for annual recommendations on agricultural price policy made by the Commission for Agricultural Costs and Prices (CACP). Presently, data are collected for 29 crops. The total number of observations in the sample varies each year and across different crops.

The sampling scheme of this survey data is a three-stage stratified random sampling. For complete coverage, the states are divided into homogeneous "zones" depending on the cropping pattern, soil type, rainfall, irrigation, etc. The sampling scheme is defined in (DACNET, 2008)

1. Tehsils are the first-stage sampling unit. These are allocated to different zones in proportion to the area under the principal crop in the zone and the total area under the crop in the state. Within a zone, the allocated number of tehsils are selected with probability proportional to the area under the principal crop with replacement.

2. A cluster of villages is the second-stage sampling unit. Within a tehsil, the cluster of villages is selected with probability proportional to the area under the principal crop in the tehsil and with replacement. A nucleus village is selected to choose a cluster of villages, and nearby villages are added. It is done such that the first village so added is the nearby village to the south of the nucleus village, and the second village is to the west of the nucleus village. Once selected for data collection under the scheme, the villages are retained for three years.

3. Operational holding within a village is the third-stage sampling unit. The holdings in the selected villages are listed in ascending order of their size and stratified into five size classes to equal the total operational area in each size class. The five size classes are defined as follows: marginal (<1 hectares), small (1-2 hectares), semi-medium (2-4 hectares), medium (4-6 hectares) and large (>6 hectares). The holdings are selected randomly from each size class to collect detailed input-output data.

From different agro-climatic zones of the states, farmers are selected using three-stage random stratified sampling. Each sample household is surveyed consecutively for three years of a block period. The fieldwork for the scheme is entrusted to the State Agricultural Universities. The Directorate of Tobacco Development is responsible for surveying tobacco.

The main variables related to labour available in CCPC data are hours of hired and family labour deployed in cultivating crops per hectare of cropland. Data on machine use, animal use, yields, and cost of labour and machine input is also collected. The agricultural operations captured in CCPC data include ploughing, sowing, manuring, irrigation, weeding, pesticide and fertilizer application, harvesting, and threshing. The survey collects information on hours spent on different tasks each day during the crop period. This information is disaggregated by family, servant, exchange and casual labour for men, women, and children. Information on daily expenditure on human labour, animal labour, and machines is recorded each week for all months.

The working of the CCPC scheme has been reviewed thrice since its implementation. S R Sen Committee (1979) recommended adopting a crop-complex approach instead of a single-crop approach. The crop-complex approach collected information on all crops in the selected sample holding. The second review was done in 1990. It recommended combining the crop-complex approach with the single-crop approach to get estimates for special/minor crops (C H Hanumantha Rao Committee, 1990). A major recommendation was that family labour, for which wages were imputed at the level of wages paid to attached farm workers, should be imputed on wages paid to the casual wage labourers. The committee also recommended that the cost of farm management be included in the estimates of total cost and be valued at ten per cent of the paid-out cost. The third review in 2005 was primarily to revise the Minimum Support Prices (MSP) estimates and tariffs for selected crops. The Review Committee recommended that the coverage of crops should be increased under this survey, and horticultural crops should be included. The committee also recommended the inclusion of marketing and transport costs in the estimates of the total cost of cultivation (Y K Alagh Committee, 2005). The Indian government accepted these recommendations with minor modifications. Sen and Bhatia (2004) pointed out that many of the recommendations of the Review Committees were not implemented.

Surjit (2017) notes that the current CCPC surveys have serious methodological problems, problems related to data processing and quality, and problems associated with the disaggregation of published data. The data released for final use is only available at farm-level disaggregation. All data recorded in the Record Type (RT) files are not released. As a result, the pattern of labour use cannot be studied for men and women separately or for different operations in crop cultivation. See Section 3 to read about the methodological challenges in CCPC data while one tries to estimate the total labour used per hectare of land.

Some gaps remain in these data that have also been criticised. Vaidyanathan (1986) argued that the FMS data did not capture post-harvest operations well. This continues to be the case with the present series of CCPC data as well. Despite these issues, Sen and Bhatia (2004) described this CCPC dataset as "a veritable goldmine of continuous farm-level data which has remained grossly underutilised".

III

METHODOLOGICAL AND DATA CHALLENGES IN USING FARM-LEVEL CCPC DATA

The cost of cultivation data has limitations, but this is the only source of information on labour used in all major crops in India. The farm-level data is available from the Directorate of Economics and Statistics website in Excel sheets from 2000 onwards. The state agricultural universities validate these data, and a part of the entire data set is released. The Record Type (RT) files, which are used to record this data at the farm level, are not validated by statistical tests and their release is limited to certain variables only. Farm-level data allows for a more nuanced study of the cost of cultivation in Indian farms. The cost of cultivation data used for price policy recommendation is the "state-averages" data released by the government as a separate dataset based on computations made on farm-level data. In principle, for the limited variables released in the state-averages data, one should be able to match the estimates using farm-level data. However, these farm-level data suffer from inconsistencies discussed as follows:

Inconsistencies in Farm-Level Observations: Computing Labour Used per Hectare or Labour Costs per Hectare

Despite advancements in data collection methods, significant challenges persist in farm-level observations recorded under CCPC data. I have divided these inconsistencies into two periods— pre-2010 and post-2010—because the nature of the issues differs significantly between these periods. Before 2010, the major inconsistencies in farm-level observations were data entry and recording errors. Post-2010, the primary challenge was the unavailability of several key variables essential for computations, which were not released or validated.

There were 5,11,957 observations at the farm level in the data from 1993 to 2010. However, state-level average estimates did not match the ones computed using farm-level data for some states in some years. Note that the method described below to correct inconsistencies can be applied to any analytical research based on CCPC farm-level data. While I used it for a study on labour use and expenditures, it can be replicated for other studies. Here is how I addressed these data inconsistencies to derive estimates of labour use and expenditure:

Drop Observations: All observations where the zonal level multiplier was infinite were dropped. All observations where the cluster factor weight was zero, or the sample size group number of growers was zero were dropped. This totalled 475 observations. There

were 567 observations where the area of the crop in the zone was given to be zero. Since this variable has to be, in principle, non-zero, these observations were also dropped.

Corrections to Match the State-level Average Estimates: After dropping these observations, estimates of labour used per hectare were computed at the state level and compared with the state-average summary data released by the Ministry. While the calculations matched for most states and years, there were 13 states for which the average labour use did not match for many observations. Upon exploring the data, I found 1,231 observations where the wrong zone code was assigned in the farm-level data. The only variable that attains the same value in observations for a particular zone for a given crop, size group, state, and year combination is the area of the crop in the zone. The mistake in these observations was of the following kind: the area belonged to, say, zone 3, but the observation was recorded as zone 2. Through a careful analysis, correct zone codes were assigned to each of these observations. However, there were basic data entry mistakes in the variable "area of crop in the zone" due to which zonal level, and therefore, state-level estimates would be incorrect. Such mistakes included jumbled-up digit errors in rounding. For some observations of a particular zone, size group, state, crop, and year combination, the area was given in round numbers, whereas, for the same combination, some observations had an area written in a non-rounded manner. Another error was that decimals were wrongly placed in the variable "area of the crop in the zone," which also required manual correction.

West Bengal, Punjab, and Haryana data were more erroneous (see Table 1).

TABLE 1. NUMBER OF OBSERVATIONS CORRECTED IN CCPC FARM-LEVEL DATA FOR DATA-ENTRY MISTAKES

State	Number of observations
West Bengal	604
Punjab	303
Haryana	174
Rajasthan	40
Madhya Pradesh	38
Uttar Pradesh	19
Assam	16
Andhra Pradesh	14
Odisha	9
Karnataka	5
Himachal Pradesh	5
Tamil Nadu	3
Maharashtra	1

Source: Calculated using farm-level CCPC data

It needs to be stated that these problems were more frequent in the data from 1993 to 2000 (978 observations) compared to data from 2000 to 2010 (253 observations). Nonetheless, correcting these mistakes did not guarantee a complete match of state-wise average estimates with the estimates from farm-level data. Though the computed numbers were much closer to the state averages, the correction was worth it. Moreover, without this correction, the zonal and state-level estimates were completely off the reported state averages and made no sense.

State-specific Issues: However, there remained cases where the computed estimates did not match the reported estimates. A major inconsistency was in the entire data series for Bajra in Gujarat, where the estimates were almost double the reported averages. To deal with this, the state of Gujarat was dropped from the analysis on bajra.

Missing Variables in Data: Another issue in the released farm-level data is that certain important variables are not released each year. In particular, the variable "size group" is not released for 2017–18. Not having these variables limits our analysis of size-class data for 2016–17 onwards. Another missing aspect is that data for the farm level for sugarcane, onion, and potato was not released each year for each state.

Issues in Calculating Farm-level Statistics: More pertinent issues must be raised here. These relate to factors on which the information is not collected or released in the CCPC dataset, which limits its use for scholars of agrarian research. These factors include total land operated and owned by the household under study. The farm-level data does not indicate how many farms belong to the same household unit. Even if we take farms as a unit of study and not households, the released data does not provide information about other crops grown on the same land. This has profound implications for studies on inputs and costs of cultivation. One cannot calculate total costs or total labour use on the farm. Moreover, the investigators collect information on labour used in each operation and by men, women, and children during the cost of cultivation surveys. But this information is also not released.

Constructing an All-India Series: Managing Gaps in State Surveys

It is crucial to carefully address outliers and gaps in the data to construct a comprehensive all-India time series. Occasionally, surveys were not conducted for certain crops in specific years within a block-year period. If a major crop-producing state was not surveyed, this omission could significantly distort India's average labour use estimates. For example, during the block year period 1981–84, surveys were conducted for sugarcane in five major states in 1981-82, seven in 1982–83, and six in 1983-84. However, Uttar Pradesh, which has the largest area under sugarcane cultivation, was not surveyed in 1983–84, though it was included in the prior two years. This omission caused a significant distortion and bias in the all-India labour use estimate for sugarcane in 1983–84. Therefore, handling such issues with care is essential to ensure accurate estimates. Analysts must be vigilant in identifying and adjusting for these gaps to maintain the integrity of the all-India series. In my research

on labour absorption, I used the following method to deal with this problem. Such years were dropped from the all-India series and the block year for the particular crop (sugarcane, for example) for any further computation. Table 2 lists all the cases where such an issue existed in the farm-level data. This issue was not included in the data collected during the 2000s.

TABLE 2. OBSERVATIONS DROPPED TO CONSTRUCT THE ALL-INDIA SERIES FOR DIFFERENT CROPS AND STATES, BY YEARS

Year	Crop	State
1974	Jowar	Maharashtra
1974	Groundnut	Gujarat
1977	Groundnut	Gujarat
1978	Rapeseed and Mustard	Assam
1978	Sugarcane	Andhra Pradesh
1980	Bajra	Gujarat
1980	Jowar	Karnataka
1981	Arhar	Madhya Pradesh
1981	Onion	Maharashtra
1981	Arhar	Madhya Pradesh
1982	Rapeseed and Mustard	Assam
1982	Rapeseed and Mustard	Uttar Pradesh
1983	Rapeseed and Mustard	Assam
1983	Rapeseed and Mustard	Rajasthan
1983	Rapeseed and Mustard	Uttar Pradesh
1990	Groundnut	Maharashtra
1990	Sunflower	Maharashtra
1992	Sunflower	Maharashtra
1993	Rapeseed and Mustard	Gujarat

IV

SITUATION ASSESSMENT SURVEY DATA

The other secondary sources of data used in this study include the Situation Assessment Survey (SAS), 2018–19, which was conducted by the National Statistical Office. These have limited data for measuring labour use, albeit these are important to obtain an aggregate picture of labour absorption. The drawback of CCPC data is that these data do not give an estimate of the aggregate level of labour use in crop cultivation. Labour use could only be ascertained for particular crops using CCPC datasets.

The Situation Assessment Survey (SAS) was first conducted in 2003 by the NSSO as part of its 59th round. It was undertaken again in 2013 as part of the 70th round and in 2019 as part of the 77th round. The SAS schedule was designed to collect

information on various aspects of farming and other socio-economic characteristics of agricultural households. The method of selecting households for these surveys differed in the three rounds. The 77th round was a major improvement as it included an integrated schedule of enquiry for the survey on Land and Livestock Holdings of Households and Situation Assessment of Agricultural Households. Another improvement was the inclusion of imputed expenses for various items of expenditure in agriculture.

Bakshi (2021) examined the SAS data in detail and pointed out that the latest round of SAS survey captured the imputed costs of production and collected detailed data on other sources of income such as pensions, rents, etc., which makes it extremely useful for estimating incomes of agricultural households. She also presented some problems in the new round, such as the lack of possibility to compute crop-wise profitability from SAS data.

For this article, I am only referring to the challenges in using these data to assess labour use in agriculture. There are two major issues. First, this survey does not include information on the number of labour days used in agriculture. Data on labour collected in this survey includes expenditure on hired labour for the farm households for the earlier rounds and the imputed cost of family labour in the recent round. Second, the expenditure on labour is not disaggregated by crop; therefore, estimates of expenditure on labour for each crop are not possible. An advantage of using this data as a proxy for labour absorption is that the CCPC data does not allow aggregate labour use estimates. Though a proxy, SAS expenditure data helps bring out the aggregate picture.

The SAS captured both casual and regular hired labour expenditure. To impute family labour expenses, the survey used the following methodology: family labour was valued at the rate of wages paid for hired labour for similar work. Men-equivalent wage rates were used to account for work done by men, women, and children where the wage rates were not available directly. They assumed that work performed by a woman was equivalent to 0.8 men's equivalent and that work done by a child was equivalent to half an adult's equivalent. The wage rate used for imputing family labour expenditure is not made public by the NSO (National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India 2021).

V

COMPARING LABOUR COSTS ESTIMATED ON THE BASIS OF SAS AND CCPC DATA

Using the SAS data, Table 3 shows the average expenditure incurred on family, hired labour, and machines per acre of operated area in different states.[2] The average expenditure on family labour for India stands at Rs 924, for hired labour at Rs 1,229 and for machines at Rs 782 per acre of the operated area. Among the major states,

[2]Expenditure incurred on repair and maintenance of machinery in the reference year is included. Expenditure of capital nature, machinery bought, and replacement of major parts of machinery is not considered in this expenditure.

Kerala and Uttarakhand had the highest expenditure on family labour. Hired labour expenses were above the all-India average for Uttarakhand, Punjab, Haryana, Uttar Pradesh, Karnataka, Kerala, Tamil Nadu, Andhra Pradesh, West Bengal, and Bihar, among the major states. The expenditure on machines was much lower than the total labour expenditure for all states.

An important point from Table 3 is that hired labour expenditure per acre is higher than family labour expenditure for all major states, barring Rajasthan and Himachal Pradesh. Given the methodology for imputing expenses on family labour in SAS, one can conclude that hired labour is much more than the use of family labour per acre of land operated in almost all major states. In Andhra Pradesh, West Bengal, Tamil Nadu, and Telangana, the expenditure on hired labour was more than twice the expenditure on family labour.

Based on SAS data, another observation on the use of family and hired labour in agriculture is that small farmers use much more family labour than hired labour. As the farm sizes rise, the expenditure on (or use of) family labour per acre also falls. The SAS data also shows that small farmers have the highest expenditure on hired labour, and as farm sizes rise, the expenditure on hired labour falls. A strikingly similar pattern is observed for expenditure on machine labour (see Table 4).

Before summarising these observations in the context of labour use in rural India, comparing the labour expenditures using the cost of cultivation of principal crops (CCPC) data is noteworthy. As this comparison cannot be made for total crops due to the lack of total crop data, we can only see expenditure on family and hired labour and machines for some major crops and major states in India in 2018–19. We chose 2018–19 as the estimates for SAS are also based on this year. Table 5 shows the per-acre expenditure on different types of labour for selected crops in selected top-producer states. The per-acre labour costs based on CCPC are much higher than those estimated from SAS data, even though the former is merely for selected crops, whereas the SAS estimates are for the entire state. As the details of the cost of inputs are better collected in CCPC data, and these serve as benchmarks for announcing the minimum prices for crops, we can conclude that SAS data highly underestimates cultivation costs. This further means that the estimates of income may be biased. A comparison of incomes and costs can further elaborate on this. However, it is beyond the scope of the study.

To summarise, SAS has two problems. One is how data on labour expenditures are collected without disaggregation at the crop level, which highly underestimates labour costs in rural India. Second, the size-classwise data analysis renders the conclusion that small farms employ the most labour, which is also incorrect as the observations for large farms were very few. The survey methodology needs major corrections, such as data on costs and the use of inputs for each crop, or at least the major crops in major states, to draw meaningful results on labour use patterns in India.

TABLE 3. AVERAGE LABOUR EXPENSES IN CROP CULTIVATION, 2018-19 (Rs./acre)

State	Family Labour	Hired Labour	Machine Labour	Total
Jammu & Kashmir	1147 (40.5%)	751 (26.5%)	931 (32.9%)	2829
Himachal Pradesh	1169 (45.4%)	1113 (43.2%)	293 (11.4%)	2575
Uttarakhand	2361 (43.7%)	2365 (43.8%)	678 (12.5%)	5404
Punjab	1330 (33.0%)	1615 (40.1%)	1082 (26.9%)	4027
Haryana	1355 (29.0%)	1939 (41.6%)	1371 (29.4%)	4665
Uttar Pradesh	1195 (32.7%)	1347 (36.8%)	1115 (30.5%)	3657
Rajasthan	600 (35.8%)	420 (25.1%)	654 (39.1%)	1674
Gujarat	566 (33.7%)	753 (44.9%)	359 (21.4%)	1678
Madhya Pradesh	556 (27.7%)	691 (34.5%)	757 (37.8%)	2004
Chhattisgarh	756 (34.8%)	873 (40.2%)	543 (25.0%)	2172
Maharashtra	808 (37.4%)	894 (41.4%)	459 (21.2%)	2161
Goa	1201 (46.6%)	1155 (44.9%)	219 (8.5%)	2575
Karnataka	1151 (39.6%)	1287 (44.3%)	466 (16.0%)	2904
Kerala	3341 (41.6%)	4550 (56.7%)	137 (1.7%)	8028
Tamil Nadu	807 (23.8%)	1765 (52.0%)	824 (24.3%)	3396
Andhra Pradesh	709 (17.5%)	2339 (57.8%)	1001 (24.7%)	4049
Telangana	727 (17.4%)	2051 (49.1%)	1403 (33.6%)	4181
Odisha	872 (31.7%)	1108 (40.2%)	774 (28.1%)	2754
Jharkhand	735 (49.2%)	453 (30.3%)	306 (20.5%)	1494
West Bengal	1502 (25.0%)	3351 (55.9%)	1144 (19.1%)	5997
Bihar	990 (26.5%)	1504 (40.3%)	1237 (33.2%)	3731
Sikkim	444 (62.1%)	264 (36.9%)	7 (1.0%)	715
Assam	1462 (46.5%)	1047 (33.3%)	632 (20.1%)	3141
Meghalaya	3130 (64.0%)	1673 (34.2%)	88 (1.8%)	4891
Tripura	2190 (38.9%)	2479 (44.0%)	966 (17.1%)	5635
Mizoram	4076 (94.7%)	227 (5.3%)	0 (0.0%)	4303
Manipur	679 (24.7%)	1448 (52.7%)	621 (22.6%)	2748
Nagaland	456 (73.2%)	131 (21.0%)	36 (5.8%)	623z
Arunachal Pradesh	541 (38.8%)	804 (57.6%)	50 (3.6%)	1395
Chandigarh	659 (14.2%)	3309 (71.1%)	687 (14.8%)	4655
Delhi	947 (25.8%)	1487 (40.6%)	1232 (33.6%)	3666
Dadra & Nagar Haveli	2217 (67.1%)	656 (19.9%)	431 (13.0%)	3304
Daman & Diu	1592 (49.2%)	1417 (43.8%)	224 (6.9%)	3233
Lakshadweep	1941 (27.6%)	5094 (72.4%)	0 (0.0%)	7035
Andaman & Nicobar Islands	1939 (65.7%)	944 (32.0%)	70 (2.4%)	2953
Pondicherry	459 (7.0%)	4628 (70.2%)	1508 (22.9%)	6595
India	924 (31.5%)	1229 (41.9%)	782 (26.6%)	2935

Source: Calculated using Situational Assessment Survey, 2018-19
Note: The value in brackets shows the percentage share of each item in the total expenditure.

TABLE 4. AVERAGE LABOUR EXPENSES, BY SIZE CLASS OF OPERATED AREA, INDIA, 2018-19 (Rs./acre)

Size Class of Operated Area	Family Labour	Hired Labour	Machine	Total
Less than 1 acre	2361 (48.3%)	1404 (28.7%)	1120 (22.9%)	4885
1 to 5 acres	1496 (38.0%)	1433 (36.4%)	1007 (25.6%)	3936
5 to 10 acres	959 (32.0%)	1214 (40.5%)	824 (27.5%)	2997
10 to 20 acres	724 (27.4%)	1192 (45.1%)	728 (27.5%)	2644
20 to 50 acres	497 (23.0%)	1063 (49.2%)	601 (27.8%)	2161
Greater than 50 acres	299 (16.3%)	1092 (59.7%)	438 (23.9%)	1829

Source: Based on data from Situational Assessment Survey, 2018-19
Note: The value in brackets shows the percentage share of each item in the total expenditure.

TABLE 5. AVERAGE LABOUR EXPENSES BASED ON CCPC DATA, BY CROP, 2018-19 (Rs./ acre)

Crop	State	Family	Hired	Machine	Total
Cotton	Andhra Pradesh	1282 (13.8%)	5948 (64.1%)	2056 (22.1%)	9286
Cotton	Gujarat	3678 (27.9%)	6627 (50.3%)	2872 (21.8%)	13177
Cotton	Haryana	6290 (50.2%)	3415 (27.3%)	2822 (22.5%)	12527
Cotton	Karnataka	2748 (29.0%)	4699 (49.6%)	2026 (21.4%)	9473
Cotton	Maharashtra	4250 (31.7%)	6686 (49.9%)	2457 (18.3%)	13393
Cotton	Punjab	2919 (20.1%)	8142 (56.1%)	3465 (23.9%)	14526
Cotton	Rajasthan	10129 (63.2%)	3346 (20.9%)	2549 (15.9%)	16024
Paddy	Andhra Pradesh	2390 (15.7%)	7657 (50.2%)	5205 (34.1%)	15252
Paddy	Gujarat	2717 (21.7%)	6405 (51.1%)	3414 (27.2%)	12536
Paddy	Haryana	3672 (31.7%)	4645 (40.1%)	3280 (28.3%)	11597
Paddy	Karnataka	4062 (31.4%)	4824 (37.3%)	4047 (31.3%)	12933
Paddy	Maharashtra	4962 (24.9%)	10197 (51.2%)	4766 (23.9%)	19925
Paddy	Punjab	2499 (23.0%)	4247 (39.1%)	4123 (37.9%)	10869
Paddy	Kerala	5268 (20.9%)	14352 (56.9%)	5612 (22.2%)	25232
Sugarcane	Andhra Pradesh	6035 (14.8%)	34033 (83.2%)	838 (2.0%)	40906
Sugarcane	Gujarat	2112 (11.0%)	14118 (73.3%)	3026 (15.7%)	19256
Sugarcane	Haryana	2409 (15.0%)	12497 (77.8%)	1165 (7.2%)	16071
Sugarcane	Karnataka	2659 (31.1%)	5422 (63.5%)	456 (5.3%)	8537
Sugarcane	Maharashtra	7735 (25.8%)	12530 (41.9%)	9660 (32.3%)	29925
Sugarcane	Punjab	10042 (34.3%)	16247 (55.4%)	3015 (10.3%)	29304
Wheat	Gujarat	2647 (33.6%)	2169 (27.5%)	3071 (38.9%)	7887
Wheat	Haryana	2657 (29.1%)	875 (9.6%)	5609 (61.4%)	9141
Wheat	Karnataka	1226 (26.5%)	1357 (29.3%)	2052 (44.3%)	4635
Wheat	Maharashtra	3538 (33.9%)	2077 (19.9%)	4812 (46.1%)	10427
Wheat	Punjab	1218 (17.2%)	1065 (15.0%)	4804 (67.8%)	7087
Wheat	Rajasthan	8790 (63.5%)	1587 (11.5%)	3456 (25.0%)	13833

Source: Based on data from Situational Assessment Survey, 2018-19
Note: Figures are per acre of operated area. Values in brackets show the percentage share of each item in the total expenditure.

VI

CONCLUSIONS

This paper examines the complexities and limitations of using large-scale survey data to assess labour use in agriculture. Data on the cost of cultivation of principal crops is available at the farm and state levels. However, due to multiple issues and the complexity of using the farm-level data, the state-level data is primarily used by scholars of agrarian research. This paper enumerates the methodological challenges of using farm-level data to compute labour used and labour costs in crop cultivation in India. The paper describes these challenges and offers solutions such that the farm-level estimates can be used for more research. In particular, a major challenge is the lack of release of data on important variables in recent years, data-entry errors, errors in validating farm-level data, not releasing data on several key variables such as the use of men, women, and children in agriculture work, labour used in livestock, and labour used across different operations in agriculture, and most importantly not identifying the farm unit due to which only crop-level estimates are possible, and not the farm-level estimates. Another source of large-scale data on agricultural labour is the situation assessment survey (SAS) data released in 2018–19. Two primary issues are identified using SAS data: the absence of data on the number of labour days used in agriculture and the lack of detailed data on crop-wise expenditure on labour. Furthermore, the labour expenditure estimated using SAS significantly underestimates agricultural labour costs compared to the CCPC data. Consequently, it leads to underestimating the total costs involved in crop cultivation.

This paper underscores the critical need to enhance data collection and dissemination methods to achieve a more precise and comprehensive understanding of labour use and expenditure in agriculture. Such improvements are essential for effective policy formulation in the agricultural sector. Validating and releasing the complete set of CCPC data, leveraging the expertise of researchers, is crucial for gaining insights into the current agrarian crisis. Despite the meticulous design of the questionnaire and the painstaking efforts in data collection, the utility of this data remains limited due to the non-release of several key variables, especially after 2017. These data must be validated and made publicly accessible.

Received: July 2024. *Revision accepted: February 2025.*

REFERENCES

Acharya, S. (1992). Labour use in Indian agriculture: Analysis at macro level for the eighties. *Indian Journal of Agricultural Economics*, 47(2), 169.

Bakshi, A. (2021). Situation assessment survey of agricultural households 2019: A statistical note. *Review of Agrarian Studies*, 11(2), 119–128.

Bardhan, P. K. (1973). Size, productivity and returns to scale: An analysis of farm level data from Indian agriculture. *Journal of Political Economy*, 81, 1370–1386.

Bhalla, S. (1987). Trends in employment in Indian agriculture, land and asset distribution. *Indian Journal of Agricultural Economics*, 42(4), 537.

Bharadwaj, K. (1974). *Production conditions in Indian agriculture: A study based on farm management surveys.* Department of Applied Economics Occasional Papers.

C. H. Hanumantha Rao Committee. (1990). Report of the Expert Committee for review of methodology of cost of production of crops. Ministry of Agriculture, Department of Agriculture & Cooperation, Government of India.

DACNET. (2008). *Manual on cost of cultivation surveys*. Central Statistical Office. Retrieved from https://eands.dacnet.nic.in/PDF/Manual%20on%20Cost%20of%20Cultivation%20Surveys.pdf.

Dantwala, M. L. (1987). Rural assets distribution and composition of labour force. *Indian Journal of Agricultural Economics*, 42(3), 275.

Gangwar, A. C. (1970). Inter-regional differences in agricultural labour use, efficiency and wages. *Indian Journal of Agricultural Economics*, 25(3), 39.

Jena, D. D. (1957). Pattern of employment of hired labour in agriculture in Orissa. *Indian Journal of Agricultural Economics*, 12(2), 28–33.

National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India. (2021). Situation assessment of agricultural households and land and livestock holdings of households in rural India, 2019. National Sample Survey Office, Ministry of Statistics & Programme Implementation, Government of India, New Delhi. Retrieved from https://pib.gov.in/newsite/printrelease.aspx?relid=166982.

Paranjpe, V. M. (1958). Input-output relations in agriculture. *Indian Journal of Agricultural Economics*, 13(1), 1–43.

Patnaik, U. (1971). Capitalist development in agriculture: A note. *Economic and Political Weekly*, 6(39), 123–130.

Prakash, H. (1962). Regional variations in labour inputs per acre. *Indian Journal of Agricultural Economics*, 17(3), 73.

S. R. Sen Committee. (1979). Special Expert Committee on Cost of Production Estimates. Government of India.

Saini, G. R. (1969). Resource-use efficiency in agriculture. *Indian Journal of Agricultural Economics*, 24(2), 1–18.

Sen, A., & Bhatia, M. S. (2004). *Cost of cultivation and farm income* (Vol. 14, State of the Indian Farmer: A Millennium Study). Academic Foundation in association with Department of Agriculture & Cooperation, Ministry of Agriculture, Government of India.

Sen, A. K. (1962). An aspect of Indian agriculture. *Economic and Political Weekly*, 14(4-6), 243–246.

Surjit, V. (2017). The evolution of farm income statistics in India: A review. *Review of Agrarian Studies,* 7(2), 39-61.

Vaidyanathan, A. (1978). Labour use in Indian agriculture: An analysis based on farm management survey data. *CDS Working Paper 72,* Trivandrum: CDS.

Vaidyanathan, A. (1986). Labour use in rural India: A study of spatial and temporal variations. *Economic and Political Weekly*, 21 (52), A130–146.

Y. K. Alagh Committee. (2005). Report of the Expert Committee to examine methodological issues in fixing MSP. Ministry of Agriculture, Department of Agriculture & Cooperation, Government of India. Retrieved from https://cacp.dacnet.nic.in/ViewQuestionare.aspx?Input=2&DocId=1&PageId=66&KeyId=514.