MID-TERM IMPACT ASSESSMENT REPORT

Odisha Particularly Vulnerable Tribal Group Empowerment and Livelihood Improvement Program (OPELIP)

Jonathan Mockshell, Sylvester Ogutu, James Garrett, Thea Ritter, Andrew Kennedy, Andrea Fongar

Alliance of Bioversity and CIAT/CGIAR

July 2022

Acknowledgements

The authors would like to thank all IFAD, PMU, and ICRISAT staff who assisted with the design and implementation of this impact assessment and provided inputs for this report. This especially includes Frew Behabtu, Romina Cavatassi, Ulac Demirag, Tisorn Songsermsawas, and Rasha Omar from IFAD. The continued support from OPELIP PMU staff is also greatly appreciated. Particularly, we would like to thank PMU Director Arthanari P. and Goutam Mohanty for their continuous support. We express our thanks to all the PMU focal points for their support during the fieldwork. We also would like to thank Swamikannu Nedumaran, Nandi Ravi, and Ravichand K. from ICRISAT-India and Abilash Ananthula (consultant) for providing technical support and coordinating the data collection. We also acknowledge the HDI consulting company hired to collect the data, specifically Braj Das and the HDI team for data collection. We thank Ricardo Labarta, Carolina Gonzalez, Edward Martey, and Collins Asante-Addo for reviewing the study materials. Excellent research assistance from Diego Alverez and Eli Akorsikumah is appreciated. Finally, we thank all the households that were interviewed for their time and patience.

This work was funded through a large grant from IFAD for the project “Linking Research to Impact: Increasing the Effectiveness of Agriculture and Food Systems in Improving Nutrition” and with additional support from the CGIAR Research Program on Agriculture for Nutrition and Health. The support of IFAD and the CGIAR is gratefully acknowledged.

Table of contents

List of abbreviations

BMI	Body Mass Index
CPI	Consumer Price Index
DID	Difference-in-Differences
FAO	Food and Agriculture Organization of the United Nations
FIES	Food Insecurity Experience Scale
GP	Gram Panchayat
HDDS	Household Dietary Diversity Score
IFAD	International Fund for Agricultural Development
MDD-Ch	Minimum Dietary Diversity for Children
MDD-M	Minimum Dietary Diversity for Men
MDD-W	Minimum Dietary Diversity for Women
MPA	Micro-project Agency
OPELIP	Odisha Particularly Vulnerable Tribal Group Empowerment and Livelihoods Improvement Program
OTELP	Odisha Tribal Empowerment and Livelihoods Project
PMU	Program Management Unit
PSM	Propensity Score Matching
PVTG	Particularly Vulnerable Tribal Group
SC	Scheduled Caste
STs	Scheduled Tribes
TLUs	Tropical Livestock Units
VDAs	Village Development Associations

Summary

Although there has been a huge investment in agricultural development interventions, few rigorous evaluations of such interventions exist and little is known about their impact, particularly among the most vulnerable communities. In recent years, governments and development agencies have also paid greater attention and made significant investments in shaping these projects to be nutrition sensitive so they contribute to improving nutrition among beneficiary households as well as the general population. Despite knowing the links between agriculture and nutrition conceptually, few evaluations of the impacts of such nutrition-sensitive projects on nutrition or other aspects of livelihoods exist.

This study reports on the findings of a midline evaluation of the effects of the Odisha Particularly Vulnerable Tribal Group Empowerment and Livelihoods Improvement Program (OPELIP) on various indicators. OPELIP is an IFAD-funded project carried out by the Government of Odisha State and the national Government of India, in partnership with micro-project agencies (MPA) and villages. OPELIP is being implemented in 12 districts of Odisha, India, and aims to reduce poverty and improve the well-being and nutritional status of about 62,356 households, including PVTGs, scheduled castes, and scheduled tribes, largely through support to improvements in agriculture-based production, marketing, and livelihoods. The midline assessment aims to provide useful mid-program implementation insights and feedback to improve ongoing program implementation and design of new aspects for the second half of the program, from 2021 to 2024, as well as for future programs.

While we expected OPELIP to have positive effects on beneficiary households, as outlined in the project design documents (e.g., increased adoption of yield-enhancing inputs and nutrient-dense crop varieties to improve dietary diversity and dietary quality; more secure land tenure; improved agricultural productivity to contribute to income growth; greater empowerment of women and an improved gender balance; and improved child nutrition and food and nutrition security), the results from a sample of about 2,000 households analyzed using difference-in-differences methods to account for potential confounding factors do not show significant effects for most of the indicators. Positive effects are only observed for use of chemical fertilizer, the value of production, monthly per adult-equivalent incomes, consumption expenditures, and poverty reduction. The insignificant effects of OPELIP on many of the outcomes defined during project design are possibly due to delays in program implementation or slow uptake of certain OPELIP interventions by beneficiaries, such as access to irrigation, adoption of yield or soil fertility improvement innovations, adoption of kitchen gardens to promote more diverse diets, land titling, and women's empowerment training. Complementary studies were also undertaken to examine these operational issues in more depth and are reported elsewhere.

The results from this quantitative evaluation show that OPELIP is contributing to improved value of production, monthly per adult equivalent incomes, consumption expenditures, and poverty reduction. However, more attention should be given to the implementation of all three program components ((1) community empowerment; (2) natural resource management, food and nutrition security, and livelihood improvement; and (3) community infrastructure and drudgery reduction) since the effects of the program on most of the indicators under the program components were insignificant. We recommend paying attention to the promotion and training on kitchen gardens, crop diversification, and nutrient-rich local crops. Increasing agricultural productivity should also be coupled with increased market access through

infrastructure development. Supporting women's participation in self-help groups and the acquisition of land titles is critical to increasing the impact of the program.

1. Introduction

The Odisha Particularly Vulnerable Tribal Group Empowerment and Livelihoods Improvement Program (OPELIP) is a comprehensive program that aims to empower and improve the livelihoods and food and nutrition security of vulnerable tribal communities in the state of Odisha (previously known as Orissa), India. Scheduled Tribes (STs) are among the poorest population groups in rural India (IFAD, 2014). Among the 705 STs, the Government of India further defines 75 as Particularly Vulnerable Tribal Groups (PVTGs). PVTGs often have a relatively small but homogenous population, are relatively more isolated physically and socially, and are even further deprived of key development indicators, such as food and nutrition security, education, and health (IFAD, 2014).

OPELIP commenced in mid-2017 and is planned to end in 2024. The program covers 12 districts in the state of Odisha and aims to improve the living standards of at least 62,356 beneficiary households, namely 32,090 PVTGs, 13,970 other ST households, 5,486 Scheduled Caste (SC) households, and 10,810 other households (IFAD, 2014). OPELIP was rolled out as a scale-up following the successful implementation of a similar IFAD-funded program called the Odisha Tribal Empowerment and Livelihoods Project (OTELP), which was implemented from 2004 to 2015 (OTELP, 2016). OPELIP uses a comprehensive approach that includes support for improved access to land and land titles, capacity building to improve gender balance, access to credit and markets, improvement of community infrastructure, and essential social services (IFAD, 2014). OPELIP is also defined by IFAD as a "nutrition-sensitive" project, meaning that its design integrated elements with the specific intent of improving the nutritional status of beneficiary households.

The midline assessment was conducted in 2021 and follows the baseline assessment of the program conducted in 2017 (the same baseline households are revisited). The midline assessment aims to provide useful mid-program implementation insights and feedback to improve ongoing program implementation and design of new aspects for the second half of the OPELIP, from 2021 to 2024. Although there has been a huge investment in agricultural development interventions, little is known about their impact, especially among vulnerable tribal groups and rural populations in South Asia. In addition, in recent years, governments and development agencies have been paying greater attention and making significant investments in shaping agricultural projects to be nutrition sensitive, so they contribute to improving nutrition among beneficiary households as well as the general population. Despite knowing the links between agriculture and nutrition conceptually, few evaluations of the impact of such nutrition-sensitive projects on nutrition or other aspects of livelihoods exist (Bird et al., 2019; Pandey et al. 2016; Ruel et al., 2018; Winters et al., 2010; World Bank, 2011).

This midline assessment contributes towards filling this knowledge gap and supports IFAD's efforts in generating a large body of evidence on the impacts of agricultural and rural development interventions on economic development and improved living standards of poor and vulnerable populations. It also adds to the impact assessment portfolio used to assess the overall impact of IFAD's projects across the globe. Generally, assessing the impact of agricultural interventions should rank high on the research and development policy agenda out of the need to generate concrete evidence for effective policymaking and/or to achieve inclusive rural transformation through agricultural development (Gertler et al., 2016).

The goal of OPELIP is to reduce poverty and improve the wellbeing of PVTGs through income growth and improved food and nutrition security for about 62,356 households, including other tribal and poor households in the intervention areas. The program aims to achieve this goal through the following four components: (1) community empowerment; (2) natural resource management, food and nutrition security, and livelihood improvement; (3) community infrastructure and drudgery reduction, and (4) program management. These four components (explained in more detail below) are expected to play a crucial role in improving the well-being of PVTG households by providing solutions to their weak community-based institutions, infrastructure, and overdependence on subsistence farming and forest gathering for their livelihoods, in addition to shedding light on the program's overall management.

This report presents the results of the midline assessment of the impacts of OPELIP on program beneficiaries using the same set of indicators in both baseline and midline survey rounds. It also includes some additional indicators, particularly concerning nutrition outcomes.

The rest of this report is organized as follows: Section 2 describes the program's theory of change (impact pathways), impact assessment questions, unintended program effects, and coverage or target population; Section 3 provides an overview of the sampling strategy and sample size across districts and villages, as well as a description of the impact assessment design and identification strategy; Section 4 presents descriptive statistics by program participation and estimation results of selected outcome indicators; Section 5 concludes.

2. Theory of change and main impact assessment questions

2.1 Development problem and theory of change

While India has achieved the status of a middle-income economy, rural poverty is still widespread. Malnutrition is also a public health concern in India with almost one-half of children below three years of age being underweight (low weight-for-age) or stunted (low height-for-age) (IFAD, 2014). Poverty varies by tribe, caste, occupation, etc. For instance, STs are among the poorest, followed by SCs. Among STs, PVTGs are the poorest and most vulnerable. India's 75 PVTGs are distributed across more than 14 states. The state of Odisha has the largest number of PVTGs (13 out of 75), with about 58,708 PVTG households (IFAD, 2020a; Rao et al., 2020). These PVTG households live in remote areas, are socially marginalized, have limited employment and income-generating opportunities, are extremely poor, and are vulnerable to the adverse effects of climate change, which affects their primary source of livelihood (subsistence farming and forest gathering). PVTGs also experience challenges with food and nutrition security, literacy, and health (IFAD, 2014). As result, there is a need to promote and support community-based institutions and infrastructure for PVTGs to reduce their vulnerability and improve their well-being.

OPELIP aims to contribute to the improved well-being of PVTGs through four program components. We briefly highlight activities, expected outcomes, and impacts of each of the four program components with the aim of showing the program's theory of change (ToC). The ToC shows the causal mechanisms (pathways) through which the program's inputs and activities are expected to lead to outcomes and impacts. According to the program's design (IFAD, 2014), the four program components are:

(1) Community empowerment. This entails the promotion of self-help groups (SHGs) primarily for rural finance and savings and village development associations (VDAs) to plan need-

based activities. Under this component, individuals through SHGs receive training on rural finance and savings. The promotion of financial skills through SHGs is expected to enable the social development of SHG members by facilitating group savings and access to credit. Under this component, the program also aims to build the capacity of VDAs by providing training to plan and execute the needs of the community, which may include natural resource management, community-based paralegal services, healthcare, nutrition education, and community infrastructure. We would expect the increased capacity from these trainings along with additional support services to enhance the formation of and membership in women's SHGs, increase savings, strengthen rural finance systems, and improve the gender balance, food and nutrition security, and health.

(2) Natural resource management, food and nutrition security, and livelihood improvement. This component entails issuing landless households with land title certificates for their households and for the land they have been cultivating. This component also aims to improve land and water resource management through the construction of irrigation and storage facilities. In addition, it aims to improve food and nutrition security through support for crop improvement activities (on-farm demonstrations, quality seed production, and farmer field schools); promotion and training on kitchen gardens, crop diversification and nutrient-rich local crops (pulses, millets, vegetables), nutritious food items (meat, eggs, fish, milk, and milk products), and fruits and spices (mango, cashew, litchi, lime, orange, turmeric, ginger, chilies, etc.); livelihood improvement activities (poultry and goat rearing); and training on the development of grain banks and local market access infrastructure. Activities under this component are expected to lead to improvements in agricultural productivity, incomes, market access, and food and nutrition security.

(3) Community infrastructure and drudgery reduction. Activities under this component include the installation of rice hullers, drying yards, milling units, and threshing floors; provision of weighing scales, household storage bins, and food processing facilities for promoting value-addition and fair trade in villages; supply of smokeless stoves and support to operations that maintain village fuel-wood reserves; and the construction and upgrading of village link roads, SHG work sheds, toilets, and gravity-fed clean drinking water supply facilities. Sufficient supply of these community infrastructure and drudgery reduction activities is expected to increase the use of clean fuel and improve sanitation and access to clean drinking water and thereby improving food security and the nutritional status of beneficiaries, as well as increasing agricultural productivity in program villages.

(4) Program management. This component involves program administration and budget management, program operations, and implementation. It aims to improve program implementation performance by ensuring that the program's goals are achieved through sound decision-making, budgeting, and effective implementation of program activities. OPELIP program management is done through the Program Management Unit (PMU) which is based in Bhubaneswar, Odisha and is responsible for the day-to-day management of the program. The PMU houses the Program Monitoring and Evaluation and Knowledge Management Unit.

Overall, we expect the program to have positive impacts on beneficiary households and villages if the interventions and services are offered as planned and if beneficiary households adopt and use the interventions and services offered. In particular, among beneficiaries, we

expect the program to lead to the adoption of improved crop varieties (improved nutrient-dense millets, pulses, cereals, fruits, and vegetables), improved agricultural productivity, income growth, income diversification, improved food and nutrition security, increased resilience to shocks, improved women's empowerment and gender balance, enhanced child nutrition and health, and increased school enrolment. At the village- or community-level, we expect the program to improve the capacity of village and community groups to identify and mobilize resources, reduce gender gaps, and lower food prices at local markets through improved access to markets and availability of food items (i.e., an enhanced local food economy). We also expect intervention villages to have better access to community resources (e.g., health centers) than comparison villages.

Figure 1 summarizes the Theory of Change for OPELIP. As mentioned, the ToC shows the causal mechanisms (pathways) through which the program's inputs and activities lead to outcomes and impacts. The ToC closely follows the program's logical framework developed by IFAD and also includes the IFAD11 indicators shown in Table A1 in the appendix. All the IFAD11 indicators are analyzed, except for indicators on youth, persons with a disability, climate change, and job creation which were not included in the baseline assessment.

Figure 1: Theory of change

Source: Modified from IFAD (2017).

2.2 Impact assessment questions

Following the ToC, the baseline assessment questions and indicators (IFAD, 2017), and the IFAD11 indicators (such as land tenure, value of production, value of market sales, household income, consumption, asset ownership, nutrition or food security, gender (e.g., women's empowerment), and resilience), the midline assessment aims to answer the following questions:

(1) Does the program through its activities, such as the provision of irrigation and seed/lentil storage infrastructure, on-farm demonstrations, and training on improved agricultural production techniques, increase the value of agricultural production or productivity?
(2) Does the program increase access to secure agricultural and homestead land tenure among PVTGs and other landless households?
(3) Does the program contribute to higher household income, per capita expenditure, asset accumulation, and reduced poverty among beneficiaries?
(4) Does the program improve child nutrition, contribute to diversified nutritious diets (pulses, milk, fish, eggs, fruits, and vegetables), and improve food security among beneficiaries?
(5) Does the program increase child school enrolment among beneficiary households?
(6) Does the program contribute to higher market sales and better market access?
(7) Does the program contribute to improvements in gender balance, women's empowerment, and intra-household decision-making?
(8) Does the program improve technology adoption and use of complementary inputs (fertilizer, pesticide, and other improved crops and agricultural practices) among program beneficiaries?
(9) Does the program increase income diversification, livelihood opportunities, and resilience to negative exogenous shocks among beneficiaries?

A large body of literature has attempted to answer these questions in different geographies and contexts, yet rigorous evidence on the impacts of agricultural development interventions on livelihood improvements of tribal communities hardly exists. While some studies have rigorously assessed the impacts of agricultural interventions on nutritional outcomes (Pandey et al., 2016; Ruel et al., 2018; Sharma et al., 2021), agricultural productivity (Davis et al., 2012; Martey et al., 2021), food and nutrition security (Banerjee et al., 2015; Fiorella et al., 2016; Jodlowski et al., 2016), women's empowerment and intra-household decision-making (Alkire et al., 2013; de Brauw et., 2014; Kafle et., 2016), poverty reduction (Cahyadi & Waibel, 2016; Ogutu et al., 2020a), asset accumulation (Dillon, 2011; Muriithi & Matz, 2015), and the adoption of improved agricultural technologies (Nakano et al., 2018; Ogutu et al., 2020b), the evidence base on the impacts of agricultural interventions on these outcomes is still thin in general and non-existent for tribal communities. In addition, the existing literature shows limited evidence of the impacts of agricultural interventions on child school enrolment, women's empowerment and intra-household decision-making, resilience, and income diversification (IFAD, 2017). Concrete evidence regarding the impact of land titling certificates, agricultural productivity, and income growth is also mixed. Lawry et al. (2017) find a positive causal impact between the land tenure system and agricultural productivity and income, but the impacts differed significantly by region. The midline assessment of OPELIP aims to address these knowledge gaps in the literature by answering the above questions.

2.3 Unintended program impacts

As mentioned, MPAs are the implementing units for OPELIP. OPELIP is designed to cover all villages within MPAs as well as all other villages outside MPA areas, but within the same Gram Panchayat (GP)[1] – a GP saturation approach (IFAD, 2017). Since OPELIP's targeting includes all households within a GP with PVTGs, the program may lead to spillover or unintended effects on non-target households. For instance, positive spillover effects may occur if program beneficiaries share agricultural knowledge received from the program with non-beneficiaries. This could lead to increased adoption of superior technologies and higher yields among both beneficiaries and non-beneficiaries. Higher yields may also result in lower local food prices and improved food security for beneficiaries and non-beneficiaries. Such spillover effects, if present, can be very difficult to measure and would lead to a downward bias (underestimation) of the estimated program effects (comparing intervention versus non-intervention villages or groups). Reliable program estimates can only be obtained by selecting a control group that is not contaminated or affected by the program (Angelucci & Di Maro, 2016).

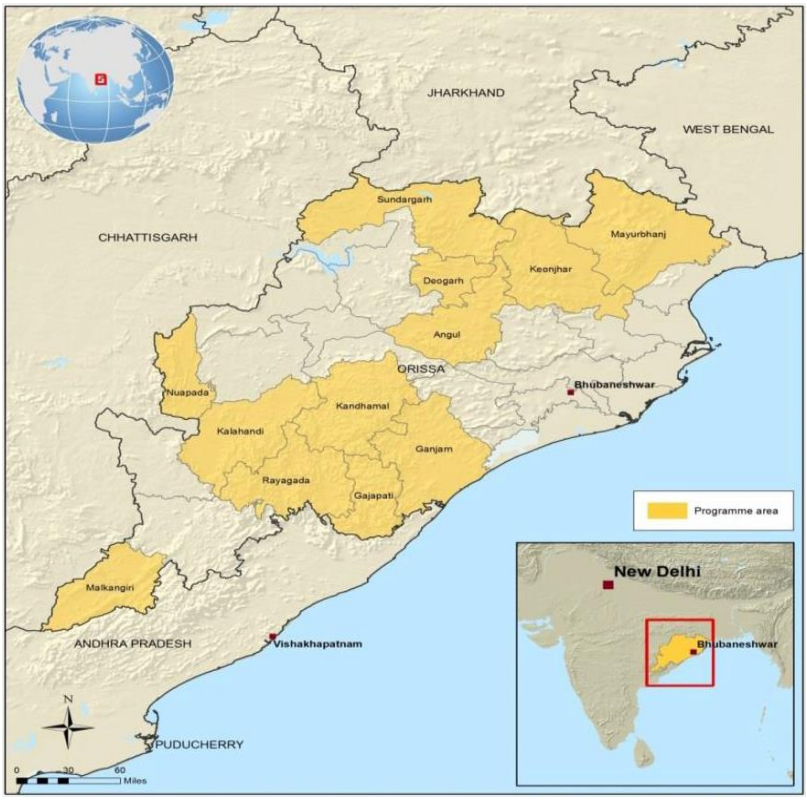
Although we expect positive impacts from the program, some undesirable impacts may also arise. For instance, increased agricultural productivity and farm incomes may increase the demand for farm labor. This may, in turn, lead to increased demand for on-farm child labor, particularly in settings where labor markets are imperfect. This implies that agricultural transformation efforts in such contexts may be achieved at the expense of children's schooling. A study by Kafle et al. (2018) showed a negative (adverse) association between increases in agricultural assets and child education through increased demand for child labor. Hence, it may be possible that children in beneficiary households may drop out of school or register poor grades as a result of increased demand for child labor.

2.4 Target population

OPELIP is implemented in 12 districts in the state of Odisha, which has the largest number of PVTGs in India. In contrast to OTELP, which targeted tribal groups, OPELIP targets the most vulnerable and ultra-poor tribal groups, who generally live in remote areas with a high prevalence of poverty (IFAD, 2014). Within the 12 districts of Odisha, OPELIP is implemented in 17 micro-project areas, which are covered by a micro-project agency (MPA). MPAs are government entities that were formed in the late 1970s. They are mandated with the responsibility of implementing special programs targeting PVTGs based on their cultural practices and beliefs. The 17 MPAs cover all 13 PVTGs in Odisha and serve as the implementing units for OPELIP. Across the 17 MPAs, OPELIP covers 84 GPs (basic village-governing or administrative units in India), 1,243 villages, and 62,356 households. The program is implemented in all villages covered by the MPAs (664 villages) and in all other villages that are outside MPA areas (579 villages) but within the same GP. There are no additional eligibility criteria, apart from the requirement to be in a GP with PVTGs. The 12 districts covered by the program are depicted in Figure 2, and the administrative hierarchy of the program coverage area is shown in Figure 3.

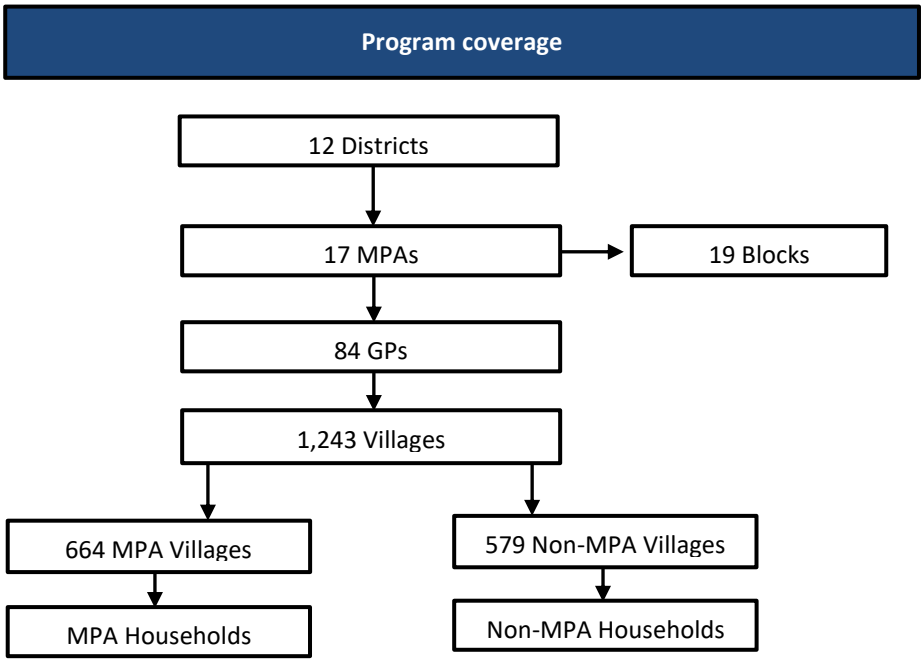
[1] A Gram Panchayat (GP), also called village council, is a local self-governing body formed by local residents.

Figure 2: The 12 districts covered by OPELIP



Source: IFAD (2014, p. vi)

Figure 3: Diagram of program coverage areas



Source: Modified from IFAD (2017).
Notes: MPA is a Micro-project agency and GP is Gram Panchayat.

3. Impact assessment design

3.1 Construction of counterfactual groups

We use the same counterfactual groups/villages that were used during the baseline survey, which was undertaken in 2017 since we are revisiting the same households (IFAD, 2017). Finding a suitable counterfactual group can be very challenging. For instance, using villages or households outside the program coverage area may be problematic as the villages may be more likely to differ from program villages in observed and unobserved characteristics. On the other hand, using households within a program area may lead to using counterfactual households that are also affected by the program. To address this challenge, the baseline survey used propensity score matching to construct counterfactual groups/villages. The approach involved finding a pool of ineligible villages outside the program areas and "matching" them on similar observed characteristics as program villages.

Using propensity score matching with the three nearest neighbors matched with replacement and the caliper length set at 0.1, counterfactual villages were drawn from outside the program area but within the same program district and Block to ensure geographical similarities between program and comparison villages.[2] This was done using 2011 census data. In particular, program and non-program villages were matched based on 20 village-level variables, which included household demographics, asset ownership, and living conditions as shown in Table 1. Due to data limitations, agricultural variables were not used. Agricultural data were only available at a higher level (the Block level) which could not be used for matching since the program is implemented at the GP level and mostly covers no more than one Block. All the variables shown in Table 1 are village-level averages.

Table 1. Village-level variables used in the propensity score matching of non-program villages

Matching variables	
Household size	Source of drinking water
Home ownership	Source of energy
Living conditions of the dwelling	Access to sanitation facility
Quality of roof materials	Use of cooking fuel
Quality of wall materials	Access to banking services
Quality of floor materials	Asset ownership (transportation)
Number of rooms in the house	Asset ownership (information)
Quality of drinking water	Bicycle ownership
Number of farming households in the village	Literacy rate
Proportion of agricultural labor in the household	Unemployment ratio

Notes: IFAD (2017). Data were obtained from a publicly available database from the 2011 India Census.

[2] A Block or Tehsil is an administrative unit that is smaller than a district, but larger than a GP.

3.2 Sample size calculations

We revisit the same households which were interviewed at the baseline survey conducted in 2017 (IFAD, 2017). Since the selection of an adequate sample size is a very crucial part of a successful impact assessment, the sample size was carefully selected at the baseline to ensure that the assessment could reliably measure the impacts of OPELIP. At the baseline, the expected change in the outcome of interest was considered more important than other factors in the sample size determination. The sample size determination strategy relied on a method developed by the World Bank that incorporates the expected minimum change in the outcome variable, its standard deviation, the critical values of the confidence interval, statistical power, and the minimum number of units to be sampled within each cluster (World Bank, 2007; Winters et al., 2010). Following this method, the sample size was determined as follows:

$$N = \frac{4\,\sigma^2(Z_\alpha + Z_\beta)^2}{D^2}[1 + \rho(m-1)]\,, \qquad (1)$$

where σ is the standard deviation of the outcome variable, Z_α is the critical value of the confidence interval, Z_β is the critical value of the statistical power, D is the minimum expected change in the average outcome variable, ρ is the intra-cluster correlation of the unit of analysis, and m is the number of units to be sampled within each cluster. The standard deviation (SD), and minimum expected change (D) are presented in Table 4. It was assumed that the analysis would have 80% statistical power and a 95% confidence level, so Z_β = 1.96, and $Z_\alpha = 1.28$. At least 15 units of observation per cluster (m) and an intra-cluster correlation of 0.05 was also assumed for the baseline sample computation (IFAD, 2017).

We summarize the details of the sample size calculation at the baseline for this impact assessment in Table 2. While there are several outcome variables of interest for OPELIP, the sample size calculation at baseline was based on only four variables – food security index, cereal productivity, rice productivity, and proportion of irrigated area – due to data limitations. Again, due to data limitations, the sample size computations were done based on district-level averages of the outcome variables rather than at the program implementation unit (GP) or village-level. Therefore, the minimum expected changes are relatively small as these changes are for the whole district. Using equation (1) and adjusting the sample size for a 10% margin of error, the largest sample size required to achieve the minimum expected change in each of the outcome variables is presented in Table 2. The largest sample size (2,096) among the four indicators was chosen to guarantee sufficient power for all outcome variables. The total sample was split equally into two to obtain a sample size of 1,048 households each for the program beneficiaries and non-beneficiary households.

Table 2. Sample size calculation

Outcome	Average of 12 districts	SD	Minimum expected change	D	N	1.1*N
Food security index	0.39	0.10	5%	0.02	1,905	2,096
Cereal productivity (kg/ha)	1,404	555.69	10%	140.4	1,118	1,230
Rice productivity (kg/ha)	1,439.58	642.13	10%	143.9	1,420	1,562
Irrigated area (%)	27.72	12.86	10%	2.8	1,537	1,691
Calculated sample size					**1,905**	**2,096**

Notes: SD is the standard deviation, D is the minimum expected change in the average of the outcome variable, and N is the sample size. The minimum expected changes are relatively small because we use district averages, but the program coverage area is only a fraction of each district. Because these expected changes are for the districts, we expect a higher level of minimum changes in the project area.
Source: IFAD (2017, p.18).

3.3 Sampling strategy

As mentioned, we interviewed the same households surveyed in 2017 in the baseline survey. The sampling strategy used for the baseline survey is as follows. At the baseline, a two-stage proportional sampling procedure was used to generate the sample. In the first stage, proportional stratified sampling was used to determine the number of households to be selected from each of the 17 MPAs covered by the program. This aimed at ensuring that there is a proportional representation of households from all the 17 MPAs and also sought to capture variations in geographical and agro-climatic conditions in each of the MPAs. Sample size calculation showed that 1,048 program households – 1.68% of 62,356 beneficiary households – would be required for the baseline. Thus, 1.68% of the households in each of the 17 MPAs covered by the program were randomly selected. In the second stage, the sample of 1,048 households to be covered by the program was divided by 12 (the minimum number of households set to be interviewed in each village) generating a total of 87 treatment villages that were randomly selected for the survey.

Details of OPELIP's sampling frame are shown in Table 3. The final sample includes 1.68% of all beneficiaries from each of the 17 MPAs to sufficiently capture differences in MPA characteristics. As mentioned, there were no MPA areas associated with the comparison group (non-program villages). Hence, 1,048 comparison/control households were selected from outside program areas, but within the same District and Block. Control households were sampled from control villages that were selected based on a propensity score matching procedure described in subsection 3.1.[3] A total of 87 villages were also sampled for the comparison group. Therefore, treatment households are drawn from 87 treated villages (villages covered by OPELIP), while control households are drawn from 87 control villages (villages not covered by the program), according to the baseline survey.

[3] From each MPA, three sets of program villages were generated by randomly selecting the required number of villages three times. Program villages in each random set were accompanied by three potential control villages as determined in subsection 3.1, using the propensity score matching approach. For each randomly selected program village, the best matched control village was identified and validated with help from PMU and MPA-level staff.

Table 3. Sampling frame for program beneficiaries: Districts, MPAs, villages, and households

District	MPA	No. of GPs in MPA	No. of Villages in MPA	No. of HHs in MPA	Final sample villages and HHs	
1. Anugul	PBDA, Jamardihi	4	43	2,482	4	42
2. Debgarh	PBDA, Rugudakudar	8	94	7,300	10	123
3. Gajapati	LSDA, Serango	3	34	1,925	3	32
	SDA, Chandragiri	10	121	5,012	7	84
4. Ganjam	TDA, Tumba	3	110	2,088	3	35
5. Kalahandi	KKDA, Lanjigarh	3	62	3,092	4	52
6. Kandhamal	KKDA, Belghar	3	63	1,568	2	26
7. Keonjhor	JDA, Gonasika	6	57	3,500	5	59
8. Malkangiri	BDA, Mudulipada	4	56	2,297	3	39
	DDA, Kudumuluguma	4	79	3,462	5	58
9. Mayuribhanj	HKMDA, Jashipur	12	156	9,907	14	167
	LDA, Morada	8	105	7,993	11	134
10. Nuapada	CBDA, Sunabeda	3	31	1,621	2	27
11. Rayagda	DKDA, Chatikona	5	124	4,637	7	78
	DKDA, Parsali	2	48	1,276	2	22
	LSDA, Puttasing	1	8	720	1	12
12. Sundargarh	PBDA, Khuntugaon	5	52	3,476	5	58
Total: 12 Districts	**17 MPAs**	**84 GPs**	**1,243 villages**	**62,356 HHs**	**88 villages**	**1,048 HHs**

Notes: There are equal number of counterfactual villages and households in each district. PBDA, Poudi Bhuyan Development Agency; LSDA, Lanjia Saora Development Agency; SDA, Saora Development Agency; TDA, Tumba Development Agency; KKDA, Kutiakandha Development Agency; JDA, Juang Development Agency; BDA, Banda Development Agency; DDA, Didayi Development Agency; HKMDA, Hill Khadia & Mankirdia Development Agency; LDA, Lodha Development Agency; CBDA, Chuktia Bhujia Development Agency; and DKDA, Dangria Kandha Development Agency.
Source: IFAD (2017, p.19).

3.4 Data collection

Two rounds of quantitative survey data were collected from smallholder farm households in Odisha, India. The baseline round was collected by the Academy of Management Studies (AMS) in 2017 before the roll-out of the program and covers 2,099 farm households. The midline round was collected by HDI in collaboration with the Alliance of Bioversity International and CIAT, and ICRISAT between October 2021 and January 2022, about four years into the program. To address possible issues of seasonality, the surveys covered the one-year period preceding the surveys in both survey rounds, so that all 12 months in a year were covered by the data. Due to sample attrition (of about 8.5% of households, evenly distributed between treatment and control groups), the follow-up survey round includes observations from 1,921 farm households. In the two survey rounds, data were collected from the sampled households on household demographics, migration, access to credit, group membership, asset ownership, agricultural production, marketing, off-farm income sources, food consumption, anthropometric indicators, and broader socioeconomic and institutional characteristics through face-to-face interviews conducted in the local language by trained enumerators. In

the midline round, additional data were collected on dietary quality indicators and adult anthropometry.

3.5 Identification strategy

This study aims to evaluate the impacts of OPELIP on key development indicators. A successful impact assessment requires a good identification strategy to identify the causal effects of the intervention by eliminating potential confounding factors. Since OPELIP's placement is not random, identification of its effects can be more challenging due to potential endogeneity arising from its purposive placement.

We use panel data which allows us to observe the study households across time and evaluate the program's effects using panel data regression models. In particular, we use the difference-in-differences (DID) estimator to account for time-invariant confounding factors (Greene, 2012). However, since we introduced some new indicators in the midline (midline) survey, we only have cross-sectional data for new indicators. Under such circumstances, the DID approach cannot be used. Hence, we use the Propensity Score Matching (PSM) to analyze the impacts of the program on the new indicators for which panel data is not available.

We now describe the DID approach, which is our main estimation approach. Assuming that Y is the outcome of interest and X is the vector of control variables, we can estimate the impact of the project (P) using the DID approach as follows:

$$Y_{it} = \beta_0 + \beta_1 Time_t + \beta_2 P_t + \beta_3 (Time_t * P_t) + \vartheta X + \alpha D + \varepsilon_{gt}, \quad (1)$$

where β_0 denotes the constant, β_1 captures the time trend or unobserved time effects, and β_2 represents the group effects or unobserved effects among program beneficiaries relative to non-beneficiaries in absence of the program. The coefficient on the interaction term β_3 is our parameter of interest, which estimates the effects of OPELIP. The DID estimator β_3 relies on the parallel or common trends assumption, which postulates that in the absence of an intervention (treatment), the difference between the 'treatment' and 'control' group is constant over time (Angrist & Pischke, 2008; Cameron & Trivedi, 2005). ϑ is the coefficient for the vector of control variables, α is the coefficient for the vector of district dummy variables interacted with time to control for possible unobserved time-varying district differences, and ε_{gt} is the error term. Subscripts g and t denote group and time, respectively.

As described above, we employ PSM to evaluate the impact of the program on new indicators, such as adult anthropometric indicators, minimum dietary diversity (MDD), MDD for women (MDD-W) and MDD for men (MDD-M). The PSM method is a non-parametric estimation technique that does not depend on the functional form or distributional assumptions. The basic idea of the PSM method is to match observations of program participants and non-participants according to the predicted probability of program participation. The PSM assumes the common support condition, which requires substantial overlap in covariates between program participants and non-participants, such that households that are being compared have a common probability of being both participants and non-participants. Hence, PSM allows for the comparison of outcomes of participants and non-participants with similar observed characteristics (for more detail, see Rosebaum and Rubin (1983)). For the purpose of robustness and comparison, we employ nearest neighbor and kernel matching methods.

4. Results

In this section, we present descriptive statistics of various outcome variables, differentiated by program beneficiary (treatment) and non-beneficiary (control) household, as well as PVTG and non-PVTG households, following the same format as the baseline evaluation. We also present and discuss the estimation results of selected outcome variables. As mentioned, the selected indicators are informed by the impact assessment questions identified in subsection 2.2, IFAD 11 indicators shown in Table A1 in the Appendix, and the baseline's Impact Assessment Report (IFAD, 2017). All results, except indicated, are based on panel data from the baseline and midline surveys. Incomes, production, sales values, and remittances reported in the midline survey were deflated to baseline (the year 2017) figures using the consumer price index (CPI) for India for the period 2017-2021 for appropriate comparison with baseline figures.[4]

4.1 Descriptive results

4.1.1 Distribution of tribal groups across the treatment and control samples

Table 4 presents the distribution of tribal groups across the treatment and control samples at baseline and midline. As mentioned, the midline sample is slightly smaller (N=1,921) than the baseline (N=2,099) sample due to attrition, which is about 8.5% overall, and distributed evenly between treatment and control groups. Results show that PVTG households, which are of particular interest, make up about one-third of the total baseline and midline samples. Larger numbers (shares) of PVTG households are found in the treatment group in both survey rounds, while there are fewer numbers (shares) of ST households. This is expected since OPELIP targets PVTG households. However, in spite of the priority given to PVTGs under this program, more than one-third of PVTG households remain outside OPELIP program coverage areas, potentially due to variation in the households' own classification of their tribal grouping from official records as reported by the baseline survey (IFAD, 2017). To prevent inconsistencies across the surveys, the same classification of the household's tribal group or caste used at baseline was used for this midline survey. The distribution of the tribal groups in our sample differs significantly from the Odisha state-level distribution because OPELIP focused on the 12 districts with the largest population of PVTGs. Overall, in both survey rounds, other STs comprise the largest share of the total sample, followed by PVTGs, SCs, and other tribes.

[4] Year 2021 figures are deflated using CPI deflator (0.824) computed using the CPI from the World Bank (2022).

Table 4. Distribution of tribal groups across treatment and control samples

Tribal group	2017		2021	
	Treatment	Control	Treatment	Control
Particularly vulnerable tribes (PVTG)	449	255	408	228
	(42.8)	(24.3)	(42.4)	(23.8)
Other Scheduled tribes (STs)	399	589	365	533
	(38.0)	(56.2)	(37.9)	(55.6)
Scheduled castes (SC)	46	71	44	67
	(4.38)	(6.77)	(4.57)	(6.99)
Other tribes	156	134	145	131
	(14.9)	(12.8)	(15.1)	(13.7)
Total	1,050	1,049	962	959
	(100)	(100)	(100)	(100)

Notes: Numbers represent the sample size. Numbers in parentheses are column percentages.

4.1.2 Household demographics

Table 5 presents summary statistics of selected variables covering household demographics of the sample households. Panel A of Table 5 shows that, on average, household heads were about 46 years old at baseline and 48 years old in the midline survey. Household heads also had about 2.5 years of education at baseline and three years of education in the midline survey, with the changes possibly occurring due to changes in the composition of household heads across the two survey rounds. About 10% and 14% of the households are headed by women at baseline and midline samples, respectively, and over 80% of household heads are married in both survey rounds. A comparison of these variables between the treatment and control groups does not reveal any significant differences between the groups in the two survey rounds.

Panel B of Table 5 presents summary statistics of household membership and composition. In both survey rounds, the results show that, on average, households have about five members, between one and two members of the household are children (less than 14 years old) or youth (15–29 years old),[5] or belong to the working-age group (30–64 years old). Households have relatively fewer members that are at least 65 years of age, with an average of 0.23 individuals in this age group per household. This is consistent with the low dependency ratio that shows that, on average, each household has about one dependent (dependency ratio of 0.6). The results also show that at least 50% of the adult population is literate, i.e., able to read and write in at least one language, in both survey rounds. The level of literacy actually increased in the period between the survey rounds (about 7 percentage points), although the difference in literacy rates between the treatment and control groups remained about the same (1.5–1.7%).

[5] This definition is according to the definition of the government of Odisha. However, it is noteworthy that according to the United Nations, “youth” consists of individuals between the age of 15 and 24 years (United Nations, 2004).

Table 5. Statistics of selected household variables by treatment status

	2017		2021	
Panel A: Household head characteristics	Treatment	Control	Treatment	Control
Age of head (years)	46.2	46.4	48.3	48.8
	(13.1)	(13.0)	(13.0)	(12.8)
Education of head (years)	2.58	2.53	3.00	2.99
	(3.89)	(3.78)	(4.01)	(3.96)
Female head (%)	10.2	9.63	13.6	13.6
	(30.3)	(29.5)	(34.3)	(34.2)
Married head (%)	86.7	85.8	82.2	82.3
	(34.0)	(34.9)	(38.3)	(38.2)
Panel B: Household characteristics				
Household size (count)	4.83	4.78	4.69	4.58
	(1.87)	(1.75)	(1.85)	(1.76)
Number of children below 14	1.49	1.43	1.36**	1.20
	(1.37)	(1.34)	(1.31)	(1.26)
Number of youth 15-29	1.40	1.39	1.33	1.38
	(1.20)	(1.18)	(1.19)	(1.23)
Number of adults 30-64	1.73	1.75	1.76	1.76
	(0.85)	(0.86)	(0.84)	(0.81)
Number of adults 65 and over	0.20	0.22	0.25	0.24
	(0.46)	(0.48)	(0.52)	(0.52)
Dependency ratio	0.66	0.65	0.63***	0.56
	(0.64)	(0.64)	(0.65)	(0.61)
Literacy rate (%)	50.9	52.4	57.2	59.0
	(30.8)	(29.5)	(30.2)	(29.0)
Observations	1,050	1,049	962	959

Notes: Point estimates are sample means. Standard deviations are in parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p$ <0.10, $^{**}p$ < 0.05, and $^{***}p$ <0.01.

4.1.3 Housing characteristics

Table 6 presents a summary of sample households' dwelling characteristics. In the baseline, almost all respondents (> 98%) owned their homes, which, consistently, had an average of about two rooms. This number increased to > 99% in the midline. A bit over half of the households' dwellings had improved roofing material, meaning they had corrugated iron sheet, concrete or cement, or a brick or asbestos roof at the baseline. This number increased to around 60 percent in the midline, with a larger share (65.6%) of households in the treatment group having improved roofing materials compared to the control group (61.6%). Around 35% of both treatment and control households had improved wall material (cement, bricks, or blocks) at baseline and around 20–25% had improved floor material (cement, tiles, or wood planks). The share of households with improved walls and improved floor materials increased substantially, to around 43–46% for both groups, by midline.

Table 6. Housing characteristics by treatment status

	2017		2021	
	Treatment	Control	Treatment	Control
Home ownership (%)	98.0*	99.0	99.9	99.8
	(14.01)	(10.19)	(3.22)	(4.56)
Number of rooms	2.32	2.33	2.22	2.14
	(1.140)	(1.044)	(1.17)	(1.12)
Improved roof material (%)	55.4	52.9	65.6*	61.6
	(49.73)	(49.94)	(47.53)	(48.65)
Improved wall material (%)	36.1	34.6	46.0	46.3
	(48.05)	(47.59)	(49.87)	(49.89)
Improved floor material (%)	25.9*	22.5	46.5	43.5
	(43.83)	(41.78)	(49.90)	(49.60)
Source of water >30 minutes, roundtrip (%)	18.3	15.7	8.73**	5.94
	(38.67)	(36.43)	(28.24)	(23.66)
Access to safe drinking water (%)	74.3	75.8	79.2***	84.0
	(43.73)	(42.86)	(40.60)	(36.64)
Access to improved sources of energy (%)	69.5***	78.2	94.0	92.6
	(46.1)	(41.3)	(23.8)	(26.2)
Access to improved toilet (%)	27.9	27.8	39.7	39.1
	(44.87)	(44.84)	(48.95)	(48.82)
Traditional cooking fuel use (%)	96.2***	97.7	90.1	91.1
	(19.15)	(14.96)	(29.85)	(28.44)
Observations	1,050	1,049	962	959

Notes: Point estimates are sample means. Standard deviations are in parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Less than 20% of households were located fairly far – more than 30 minutes (round trip) – from a source of water at baseline, while the share dropped to less than 10% at midline. This may suggest that the large majority of households now have relatively good access to water. Access improved by about 10 percentage points for both treatment and control groups and so the control group continued to have slightly better access than the treatment group. This is unexpected given that under the community empowerment component, OPELIP aims to improve access to drinking water and sanitation by connecting households through overhead tanks, either filled through gravity where a water source is available upstream or by pumping water from a tube well or stream if electricity is available (IFAD, 2014). Overall, more than 75% of the households have access to safe drinking water, which suggests that access to water is not a major problem, even though a smaller share of treatment households has access to clean drinking water compared to control households in the midline sample.

The results of Table 6 also show that more than two-thirds of treatment households have access to improved sources of energy (e.g., solar, electricity) for home use (lighting and cooking) at baseline, as do more than three-quarters of control households. In the midline survey, access to improved sources of energy among both groups increased, with treatment households' access now exceeding that of control households (94% to 93%, respectively). At

the same time, although many households have access to improved sources of energy, more than 90% of all households still use traditional sources of energy, such as wood and dung, although this has declined from around 96–97% of households at baseline. Lastly, only about 39% of the households in both groups have access to improved toilets, although this has increased from about 28% for both groups at baseline. This points to some improvement, but also to continuing sanitation challenges. Overall, a comparison of the households' dwelling conditions at baseline and midline shows improvements in access to water, quality of materials of roofs, walls, and floors, as well as access to improved energy sources and toilets, although generally, the improvements are similar for both treatment and control households, with the exceptions noted above for access to safe drinking water, roofing material and improved energy sources.

4.1.4 Schooling outcomes

Table 7 shows summary statistics of schooling outcomes among sample households by program participation. The outcomes mainly focus on the highest level of education attained by the household head and school-aged boys and girls currently attending school. On average, household heads have very low levels of education, about 2.5 years of education at baseline and about 3 years in the midline survey.

Table 7. School enrolment by treatment status

	2017		2021	
	Treatment	Control	Treatment	Control
Household level outcomes				
Highest grade completed by head	2.58	2.53	3.00	2.99
	(3.89)	(3.78)	(4.01)	(3.96)
Number of school-age children currently attending school	1.02	1.08	0.80	0.78
	(1.17)	(1.15)	(1.08)	(1.06)
Proportion of children (aged 5–18) attending school (%)	0.63**	0.67	0.54*	0.59
	(0.40)	(0.39)	(0.44)	(0.43)
Share of boys currently attending school (%)	0.65**	0.71	0.53**	0.59
	(0.43)	(0.41)	(0.47)	(0.47)
Share of girls currently attending school (%)	0.63	0.66	0.57	0.59
	(0.44)	(0.43)	(0.47)	(0.46)
Observations	1,050	1,049	962	959

Notes: Point estimates are sample means. Standard deviations are in parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p< 0.10$, $^{**}p< 0.05$, and $^{***}p<0.01$.

Results of Table 7 also show that on average, a household has one child attending school, with between 54–67% of school-aged children attending school, with the share of boys declining proportionately more than that of girls. A significantly smaller share of children in treatment households is attending school compared to control households in the two survey rounds, with a lower share of boys than girls attending school among treatment households.

Apart from comparing treatment and control households, we also compare schooling outcomes by PVTG and non-PVTG households as shown in Table 8. In both survey rounds, PVTG households have significantly lower household head education and a lower share of school-aged children (both boys and girls) attending school compared to non-PVTG households. As mentioned, improvements in the household head's education in the midline survey could be due to changes in the composition of the household head as we do not expect the program to support a return of adults to the school, but enrollment of children in school could definitely change. Overall, shares of school-attending children seem smaller in the midline sample compared to the baseline sample, perhaps due to the effects of the COVID-19 pandemic.

Table 8. Schooling outcomes by tribal group

	2017		2021	
	PVTG	Non-PVTG	PVTG	Non-PVTG
Household level outcomes				
Highest grade completed by head	1.46***	3.15	2.28***	3.36
	(2.89)	(4.15)	(3.51)	(4.16)
Number of school-age children currently attending school	1.00	1.07	0.85	0.76
	(1.14)	(1.17)	(1.10)	(1.06)
Proportion of children (aged 5–18) attending school (%)	0.57***	0.70	0.54*	0.58
	(0.41)	(0.38)	(0.44)	(0.44)
Share of boys currently attending school (%)	0.58***	0.74	0.52**	0.58
	(0.45)	(0.40)	(0.47)	(0.47)
Share of girls currently attending school (%)	0.57***	0.69	0.56	0.59
	(0.44)	(0.43)	(0.46)	(0.47)
Observations	704	1,395	636	1,285

Notes: Point estimates are sample means. Standard deviations are in parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.1.5 Migration and remittances

PVTGs are typically characterized by a lifestyle of migration in search of better farmland and livelihoods (IFAD, 2017). Against this background, Table 9 shows a summary of migration and remittance statistics by treatment status. Migration entails both local and regional movement of individuals out of the household for more than one continuous month in the last 12 months before each survey round rather than temporary movement out of the household to hunt and gather food. Panel A shows that on, average, about 27% and 24% of households have a migrant at baseline and midline samples, respectively. At baseline, the households have an average of one migrant, but in the midline survey, the number increases to five migrants among households with migrants. Around 40% of the households received remittances ranging between 9,500 to 14,500 rupees/year (panel B), on average. Although the number of migrants is higher in the midline survey, it has not contributed to increased remittances to the households, possibly due to the COVID-19 pandemic, which significantly reduced jobs and economic opportunities for migrant workers across India's major cities (Sahoo et al., 2022). A comparison of the migration and remittance variables between treatment and control groups

does not reveal any significant differences between the groups in the baseline and midline samples.

Table 9. Migration and remittances by treatment

	2017		2021	
Panel A: Migration outcomes	Treatment	Control	Treatment	Control
Household has a migrant (%)	27.2	27.4	23.8	24.3
	(44.5)	(44.6)	(42.6)	(42.9)
Observations	1,050	1,049	962	959
If household has a migrant …				
Number of migrants	1.34	1.34	5.22	5.15
	(0.63)	(0.63)	(1.74)	(1.70)
Panel B: Remittance outcomes				
Received remittances last 12 months (%)	42.0	40.8	41.5	44.2
	(49.4)	(49.2)	(49.4)	(49.8)
Amount of remittances (rupees/year)	14,474	12,042	9,552.9	11,158.9
	(36,926)	(22,670)	(19,906.2)	(31,236.0)
Observations	286	287	229	233

Notes: Point estimates are sample means. Standard deviations are in parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. 1 USD = 65.1 INR (Indian rupees) in 2017, the reference/base year.[6]

Table 10 compares migration and remittance statistics between PVTG and non-PVTG households. About 27% of PVTG households have a migrant at baseline and 24% at the midline. Out of the households with migrants, the average number of migrants is 1.34 at baseline, but the number increases to five in the midline sample, perhaps due to COVID-19 as mentioned above. The share of PVTG households with a migrant and the number of their migrants do not differ significantly from those of non-PVTG households in the two survey rounds. About 40% of the sample received remittances, ranging between 10,000 to 15,000 rupees/year, depending on the sub-sample. As shown, PVTG and non-PVTG households do not differ in their migration and remittance characteristics in the two survey periods, except for the amount of remittances received at baseline, which is significantly lower for PVTG households compared to non-PVTG households.

[6] https://www.exchangerates.org.uk/USD-INR-spot-exchange-rates-history-2017.html

Table 10. Migration and remittances across tribal groups

	2017		2021	
Migration outcomes	PVTG	Non-PVTG	PVTG	Non-PVTG
Household has a migrant (%)	27.9	26.9	23.3	24.4
	(44.9)	(44.4)	(42.3)	(42.9)
Observations	704	1,395	636	1,285
If household has a migrant…				
Number of migrants	1.33	1.34	5.31	5.12
	(0.60)	(0.64)	(1.78)	(1.70)
Received remittances last 12 months (%)	38.6	42.8	42.6	43.0
	(48.8)	(49.5)	(49.6)	(49.6)
Amount of remittances (rupees/year)	9,980*	14,972	9,876	10,592
	(19,406)	(35,001)	(36,530)	(19,644)
Observations	197	376	148	314

Notes: Point estimates are sample means. Standard deviations are in parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. 1 USD = 65.1 INR (Indian rupees) in 2017, the reference/base year.

4.1.6 Access to credit

We measure access to credit using outcomes for loan application, approval, amount, and use. Under the community empowerment component, OPELIP aims to promote rural finance and savings through SHGs where individuals receive training on rural finance and saving. Hence, SHGs could play a crucial role in promoting access to credit among treatment households. Table 11 presents summary statistics for access to credit by program participation at baseline and midline.

Results of Table 11 show that a significantly smaller share of treatment households applied for a loan compared to control households at baseline (17% versus 20%, respectively). The initial low rates of loan application observed are expected because most of the sample households are located in remote rural areas with limited access to financial institutions. In the midline survey, the share of households that applied for loans does not differ significantly between treatment and control households. However, we see that the percentage of treatment households that applied for a loan increased by over 5 percentage points, versus a bit over 1 percentage point for the control group.

While about one of every five households in the sample applied for a loan, the approval rates of applications were quite high. In fact, almost all loans were approved. Furthermore, while over 90% of applicants were approved even at baseline (94% of treatment applicants and 97% of control applicants), these numbers increased to 98% for both groups at midline. Loan amounts averaged about 17,000 rupees per year, for both groups at baseline, but increased to about 18,000 rupees for the treatment households and 19,000 rupees for the control households. SHGs were the primary source of loans for about 30% of treatment households and 20% of control households at baseline, but accounted for well over 50% of loans in the midline sample (59% and 55% for the treatment and control households, respectively), suggesting a growing role of SHGs in the provision of loans. A significantly larger share of the treatment group (33%) used loans for consumption expenditures compared to the control

Table 11. Access to credit by treatment status

	2017		2021	
In the last 12 months	Treatment	Control	Treatment	Control
Household applied for loan (%)	16.6**	20.0	22.3	21.4
	(37.2)	(40.0)	(41.6)	(41.0)
Observations	1050	1049	962	959
If applied for loan…				
Loan approved (%)	94.3	96.7	98.1	98.0
	(23.3)	(18.0)	(13.6)	(13.9)
Amount of loan (rupees/year)	17,311	16,505	17,839	19,080
	(18,736)	(14,773)	(19,097)	(18,529)
Primary source of loan (1 = SHG†, 0 = other)	28.2	21.4	59.3	55.1
	(45.1)	(41.1)	(49.2)	(49.9)
Primary use of loan (1 = consumption, 0 = other)	33.3**	22.9	11.7	8.29
	(47.3)	(42.1)	(32.2)	(27.6)
Observations	174	210	214	205

Notes: Point estimates are sample means. Standard deviations are in parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. †SHG stands for self-help groups. 1 USD = 65.1 INR (Indian rupees) in 2017, the reference/base year.

group (23%) at baseline, but the difference was not significant in the midline sample (12% versus 8%, respectively). Use of loans for consumption expenditures fell quite significantly for both groups at midline. The share of the households that applied for loans, the percent of loans approved, the amounts received, and the primary source of loans do not differ significantly between treatment and control households in both survey rounds.

Table 12 presents summary statistics of access to credit differentiated by tribal groups. The differences in loan application rates between PVTG and non-PVTG households are insignificant at baseline, but in the midline survey, a smaller share of PVTG households applied for loans compared to non-PVTG households. A similar pattern is also observed with loan approval rates, where at baseline, the loan approval rates between PVTGs and non-PVTGs are statistically insignificant, but in the midline survey, a smaller share of PVTG households received loan approvals compared to non-PVTG households. These midline sample results are quite unexpected since part of OPELIP's objectives was to strengthen SHGs and promote rural finance and savings among the PVTGs. Table 12 also shows that the amount of loan approved does not differ in both baseline and midline samples, a smaller share of PVTG households rely on SHGs as their primary source of credit compared to non-PVTG households in both survey rounds, while PVTG and non-PVTG households do not differ in their usage of loans.

Table 12. Access to credit by tribal groups

	2017		2021	
In the last 12 months	PVTG	non-PVTG	PVTG	non-PVTG
Household applied for loan (%)	19.0	17.9	17.6***	23.9
	(39.3)	(38.4)	(38.1)	(42.6)
Observations	704	1,395	636	1,285
If applied for loan…				
Loan approved (%)	94.8	96.0	95.5**	99.0
	(22.3)	(19.6)	(20.7)	(9.85)
Amount of loan (rupees/year)	17,648	16,452	17,539.5	18,776.4
	(19,866)	(14,698)	(19,431.8)	(18,598.0)
Primary source of loan (1 = SHG†, 0 = else)	11.9***	31.2	48.2**	60.6
	(32.5)	(46.4)	(50.2)	(48.9)
Primary use of loan (1 = consumption, 0 = other)	28.4	27.2	13.4	8.79
	(45.2)	(44.6)	(34.2)	(28.4)
Observations	134	250	112	307

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. †SHG stands for self-help groups. 1 USD = 65.1 INR (Indian rupees) in 2017, the reference/base year.

4.1.7 Land characteristics and farming practices

Under the natural resource management, and food and nutrition security component, OPELIP aims to improve food and nutrition security through various activities that include improvement of land tenure and support for crop improvement. Against that background, Table 13 presents a summary of statistics on land characteristics and farming practices. Panel A shows that at least 75% of the households owned land in both survey rounds, with some increase in landownership by midline. Less than 1% of households leased land in both survey rounds. Women own at least one plot of land in about half of the households in both survey rounds, but the share dropped to 40–41% for both groups at midline. About 17% of the households had land titles at baseline, but in the midline sample, this share more than doubled for both groups (34–37%). Panel A of Table 13 also shows that at least 90% of households cultivated their land – about 1.5 acres at baseline and 2 acres in the midline sample – in the last 12 months. Comparison of the variables in panel A between the treatment and control groups does not reveal any significant differences between groups in either survey round.

Table 13. Access to land, land holding size, and farming practices by treatment

	2017		2021	
Land characteristics: Panel A	Treatment	Control	Treatment	Control
Land ownership (%)	76.1	76.9	81.0	82.6
	(42.7)	(42.1)	(39.3)	(37.9)
Household leases land (%)	0.48	0.38	0.21	0.00
	(6.89)	(6.17)	(4.56)	(0.00)
Share of women owning land (%)	49.0	48.6	40.9	39.6
	(50.0)	(50.0)	(49.2)	(48.9)
Land title, *Patta* (%)	17.2	16.1	33.9	36.5
	(37.8)	(36.8)	(47.4)	(48.2)
Crop cultivation in the last 12 months (%)	92.2	92.2	90.2	90.3
	(26.8)	(26.9)	(29.7)	(29.6)
Land holding size (acres)	1.57	1.42	2.05	2.06
	(2.20)	(2.34)	(2.49)	(4.64)
Farming practices: Panel B				
Slash and burn (%)	8.38	6.58	3.74	2.61
	(27.7)	(24.8)	(19.0)	(15.9)
Contour farming (%)	13.5	11.9	3.53**	1.77
	(34.2)	(32.4)	(18.5)	(13.2)
Regular/settled agriculture (%)	76.9	78.2	86.9*	89.5
	(42.2)	(41.3)	(33.8)	(30.7)
Number of crops grown (count)	0.85	0.88	1.10	1.15
	(0.35)	(0.33)	(0.75)	(0.80)
Household irrigates a plot (%)	7.81**	5.53	45.4	46.3
	(26.8)	(22.9)	(49.8)	(49.9)
Observations	1,050	1,049	962	959
Plot-level input use: Panel C				
High yielding seed used on plot (%)	10.5**	7.80	34.7	35.4
	(30.7)	(26.8)	(47.6)	(47.8)
Chemical fertilizer used on plot (%)	44.4*	48.4	75.0*	72.1
	(49.7)	(50.0)	(43.3)	(44.8)
Manure used on plot (%)	35.8	36.5	13.7***	17.9
	(48.0)	(48.2)	(34.4)	(38.3)
Herbicides/pesticides used on plot (%)	17.1*	13.6	18.2*	16.1
	(37.6)	(34.3)	(38.6)	(36.7)
Observations	1,508	1,418	2,223	2,242

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Panel B of Table 13 shows a summary of the farming practices households were engaged in during the 12 preceding months. Less than 10% of farm households in both groups practiced slash and burn or shifting cultivation at baseline, while about 10% practiced contour farming. Shares of both practices fell to less than 5% in the midline survey, with a fairly substantial drop in use of contour farming (around a 10 percentage point drop in both groups). Over three-quarters of households in both groups practiced regular agriculture (settled agriculture) at baseline, with a slight increase to 87–90% in the midline sample.

On average, households cultivate one crop, suggesting that sample farms are less diversified in their production. Less than 10% of the households irrigated at least one of their plots at baseline, with a larger share of treated households practicing irrigation compared to control households. In the midline sample, the share of households that irrigated a plot increased to at least 45% for both groups, but the differences between treatment and control households are statistically insignificant. Except for contour farming and regular agriculture, the differences in the use of other farming practices presented in panel B of Table 13 are statistically insignificant in the midline survey. Part of OPELIP's objectives, under Component 2, is to improve food and nutrition security through support for crop diversification. However, the low average number of crops grown by sample households in the midline survey may imply that this objective has not yet been realized.

In panel C of Table 13, we show the results of input use at the plot level. High yielding seeds were used on around 11% of the plots at baseline by the treatment groups and used somewhat less, 8%, for the control group. At midline, the two groups were about the same, at around 35%. Fertilizer was less likely to be applied on treated farm household plots compared to the control households' plots at baseline (44% and 48%, respectively), but these numbers increased greatly by midline for both groups, with treatment groups now exceeding the control group (75% and 72%). Manure was used in around 35% of the plots of both groups at baseline, but in the midline sample manure use fell substantially for both groups (from 36% to 14% for the treatment group and from 37% to 16% for the control group). In both survey rounds, herbicides and pesticides were more likely used in treatment group plots than control group plots (17% versus 14% in 2017 and 18% versus 16% in 2021).

Table 14 presents summary statistics of land characteristics and farming practices differentiated between PVTG and non-PVTG households. At baseline, a significantly smaller share of the PVTG households compared to non-PVTGs own land, lease land, have a land title, and cultivated crops in the 12 months prior to the survey. However, in the midline sample, differences between PVTG and non-PVTG households with respect to these variables are statistically insignificant, suggesting an improvement in the indicators, especially among PVTG households in the midline survey given that they were worse-off at baseline.

Table 14. Access to land, land holding size, and farming practices by tribal group

	2017		2021	
Land characteristics	PVTG	Non-PVTG	PVTG	Non-PVTG
Land ownership (%)	74.0*	77.8	80.0	82.6
	(43.9)	(41.6)	(40.0)	(37.9)
Household leases land (%)	0.0**	0.65	0.16	0.078
	(0.0)	(8.01)	(3.97)	(2.79)
Share of women owning land (%)	49.9	48.3	36.8	41.9
	(50.0)	(49.9)	(48.3)	(49.4)
Land title, *Patta* (%)	13.5***	18.3	37.1	34.2
	(34.2)	(38.7)	(48.3)	(47.5)
Crop cultivation in the last 12 months (%)	90.5**	93.0	91.5	89.6
	(29.4)	(25.4)	(27.9)	(30.5)
Land holding size (acres)	1.60	1.44	1.91	2.13
	(2.87)	(1.91)	(2.37)	(4.23)
Farming practices: Panel B				
Slash and burn (%)	13.1***	4.66	3.46	3.04
	(33.7)	(21.1)	(18.3)	(17.2)
Contour farming (%)	21.0***	8.53	3.93**	2.02
	(40.8)	(27.9)	(19.4)	(14.1)
Regular/settled agriculture (%)	66.8***	82.9	87.3	88.6
	(47.1)	(37.6)	(33.4)	(31.7)
Number of crops grown (count)	0.85	0.87	1.12	1.13
	(0.36)	(0.34)	(0.71)	(0.72)
Plot irrigated (%)	8.95***	5.52	56.6***	40.5
	(28.6)	(22.8)	(49.6)	(49.1)
Observations	704	1,395	636	1,285
Plot-level input use: Panel C				
High yielding seed used on plot (%)	7.31**	10.1	31.6***	36.6
	(26.1)	(30.2)	(46.5)	(48.2)
Fertilizer used on plot (%)	25.1***	57.2	59.3***	79.8
	(43.4)	(49.5)	(49.1)	(40.2)
Manure used on plot (%)	44.3***	32.0	23.0***	12.6
	(49.7)	(46.7)	(42.1)	(33.2)
Herbicides/pesticides used on plot (%)	13.8	16.1	16.0	17.6
	(34.5)	(36.8)	(36.7)	(38.1)
Observations	1,002	1,924	1,399	3,066

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < .01$.

Panel B of Table 14 presents a summary of the farming practices sample households were engaged in during the 12 months before the survey. At the baseline, significantly larger shares of PVTG households compared to non-PVTG households practiced slash and burn (shifting cultivation), contour farming, or had irrigated a plot. However, a significantly smaller share of PVTG households practiced regular/settled agriculture compared to non-PVTG households at baseline. These results are plausible given that PVTG households tend to practice shifting cultivation and less regular/settled agriculture compared to non-PVTG households (IFAD 2014; IFAD, 2017). In the midline sample, significantly larger shares of PVTG households compared to non-PVTG households also practice contour farming and irrigation. The share of households with irrigated plots increased significantly from less than 10% to over 40% in the midline sample.

Lastly, panel C of Table 14 shows a comparison of input use at the plot level between PVTG and non-PVTG households. As shown, high yielding seeds and fertilizer were less likely to be applied on PVTG household plots compared to non-PVTG household plots in both survey rounds. For both groups, herbicides and pesticides were used in 14–18% of the plots in both survey rounds. However, larger shares of PVTG household plots had manure applied compared to non-PVTG household plots in both survey rounds.

4.1.8 Women's empowerment

Under the community empowerment component, OPELIP aims to improve women's empowerment and gender balance through the promotion of women SHGs, gender training, and rural finance services to enable the social development of SHG members. We use the same women's empowerment index used in the baseline survey, which is comprised of a total of 23 intra-household decision-making questions that capture the dynamics of decision making in seven decision domains (IFAD, 2017). The decision domains are agricultural input use, use of cash income, crop and livestock sales, food items' purchase, non-food items' purchase, child school enrolment, and another domain which includes questions related to wage employment, group membership, and loan application. Women's empowerment is measured using the proportion of women's independent decisions over the total decisions made in the household, with all the decisions given equal weight (IFAD, 2017). We also compute the share of men's independent and joint decisions over the total decisions made in the household for comparison.

Table 15 presents a summary of the women's empowerment measures by treatment status. Panel A shows that at baseline for both treatment and control groups about 10% of the household decisions were made independently by women, at least 85% of household decisions were made independently by men, and less than 3% of decisions were made jointly. In the midline sample, the share of solitary decisions made by women remains the same, but the proportion of decisions made independently by men decreased to about 63%, while the share of joint decisions increased substantially to about 24%. A breakdown of the share of solitary decisions made independently by women across the seven decision domains as shown in Panel B of Table 15 reveals the same decision-making trend observed when all decisions are aggregated. In all the decision domains, women make around 11% of the decisions independently and significant differences are not observed across treatment and control households in both survey rounds. Overall, these results show that men dominate decision-making and resource-control among the sample households.

Table 15. Proportions of women's and men's decisions by treatment

Decisions	2017		2021	
Panel A: Decision-making in all seven decision domains combined	Treatment	Control	Treatment	Control
Women's solitary decisions	0.101	0.095	0.109	0.114
	(0.301)	(0.293)	(0.308)	(0.314)
Men's solitary decisions	0.876	0.885	0.625	0.632
	(0.327)	(0.318)	(0.452)	(0.459)
Joint decisions	0.022	0.020	0.237	0.236
	(0.141)	(0.137)	(0.389)	(0.399)
Panel B: Share of solitary decisions by women in each of the seven decision domains	Treatment	Control	Treatment	Control
Agricultural input use	0.101	0.095	0.113	0.117
	(0.302)	(0.293)	(0.312)	(0.317)
Cash income use	0.104	0.085	0.118	0.115
	(0.304)	(0.273)	(0.312)	(0.310)
Crop and/or livestock sales	0.100	0.093	0.111	0.107
	(0.298)	(0.288)	(0.302)	(0.296)
Food purchase	0.101	0.096	0.122	0.117
	(0.301)	(0.295)	(0.326)	(0.321)
Non-food purchase	0.101	0.093	0.113	0.107
	(0.301)	(0.289)	(0.316)	(0.309)
Child schooling	0.101	0.094	0.112	0.112
	(0.301)	(0.292)	(0.316)	(0.315)
Other decisions	0.101	0.092	0.119	0.112
	(0.301)	(0.287)	(0.313)	(0.306)
Observations	1,050	1,049	962	959

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 16 presents a summary of empowerment measures differentiated by PVTG and non-PVTG households. The results are consistent with those of Table 15 which show that decision-making and resources are largely controlled by men. However, the results in Table 16 show some significant differences. Panel A shows that at baseline, the share of joint decisions made by PVTG households is significantly lower than the share of joint decisions made by non-PVTG households. In the midline sample, a larger share of women made solitary decisions in PVTG households compared to non-PVTG households, but the share of joint decisions made by PVTG households was lower than that of non-PVTG households. Consistent with the aggregated results in Panel A, a breakdown of the share of decisions made independently by women across the seven decision domains shown in Panel B of Table 15 reveals that a significantly larger share of women made solitary decisions in all the seven decisions domains in PVTG households compared to non-PVTG households. Since the share of solitary

decisions made by women is our main interest, improvements in the variable among PVTG households are certainly welcome.

Table 16. Proportions of women's and men's decisions by tribal group

Decisions	2017		2021	
Panel A: Decision-making in all seven decision domains combined	PVTGs	non-PVTGs	PVTGs	non-PVTGs
Women's solitary decisions	0.112	0.091	0.131*	0.103
	(0.316)	(0.287)	(0.332)	(0.300)
Men's solitary decisions	0.880	0.881	0.625	0.630
	(0.325)	(0.321)	(0.457)	(0.455)
Joint decisions	0.008***	0.028	0.214*	0.248
	(0.086)	(0.158)	(0.376)	(0.402)
Panel B: Share of solitary decisions by women in each of the seven decision domains	PVTGs	non-PVTGs	PVTGs	non-PVTGs
Agricultural input use	0.112	0.091	0.135*	0.105
	(0.316)	(0.287)	(0.336)	(0.303)
Cash income use	0.104	0.088	0.141**	0.105
	(0.303)	(0.279)	(0.340)	(0.296)
Crop and/or livestock sales	0.111	0.089	0.132**	0.098
	(0.312)	(0.283)	(0.324)	(0.286)
Food purchase	0.112	0.092	0.151***	0.104
	(0.316)	(0.289)	(0.357)	(0.304)
Non-food purchase	0.114*	0.088	0.141***	0.095
	(0.316)	(0.283)	(0.346)	(0.293)
Child schooling	0.112	0.090	0.138***	0.099
	(0.316)	(0.287)	(0.346)	(0.299)
Other decisions	0.113*	0.088	0.142***	0.102
	(0.316)	(0.281)	(0.340)	(0.292)
Observations	704	1,395	636	1,285

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.1.9 Agricultural yield and sales

Table 17 presents summary statistics on the total value of production (including both crops and livestock) in rupees; selected crop yields in kilograms per hectare (kg/ha), value of sales in rupees, and crop sales in kg/ha by treatment status. Panel A shows that, on average, sample households produced output worth 48,000 rupees at baseline and 64,000 rupees at the midline survey during the Kharif season.[7]

[7] Following the baseline analysis, we report results of Kharif season for proper comparison and to create panel data. Baseline data only covered the Kharif season; however, the midline survey includes Kharif, Rabi, and summer seasons.

Table 17. Agricultural yield and sales in Kharif season by treatment

	2017		2021	
	Treatment	Control	Treatment	Control
Panel A: Produced output				
Household-level				
Value of total production (rupees)	46,798.2	49,054.2	63,598.6	63,592.8
	(50,659.6)	(55,372.6)	(78,753.9)	(81,622.2)
	[1,050]	[1,049]	[962]	[959]
Plot-level				
Paddy yield (kg/ha)	1,530.2	1,544.7	2,546.3	2,598.8
	(1,409.2)	(1,419.7)	(1,143.8)	(1,227.6)
	[691]	[710]	[1,383]	[1,499]
Maize yield (kg/ha)	1617.0	1,482.5	2,418.0	2,382.9
	(1577.6)	(1,303.8)	(1,321.5)	(1,271.7)
	[85]	[68]	[171]	[121]
Mandia (finger millet) yield (kg/ha)	1,046.4	939.9	439.3	452.0
	(1,296.8)	(1,170.4)	(703.8)	(629.6)
	[97]	[117]	[72]	[61]
Panel B: Sold output				
Household-level				
Value of total sales (rupees)	3,821.5	3,474.4	11,278.4	11,100.0
	(9,030.2)	(8,594.5)	(19,845.7)	(19,903.5)
	[1,050]	[1,049]	[962]	[959]
Plot-level				
Paddy sold (kg/ha)	159.5	126.0	126.0	124.9
	(600.4)	(582.7)	(553.9)	(578.7)
	[691]	[710]	[1,383]	[1,499]
Maize sold (kg/ha)	180.8	216.2	1511.7**	1,123.8
	(886.4)	(821.9)	(1,619.1)	(1,571.7)
	[85]	[68]	[171]	[121]
Mandia (finger millet) sold (kg/ha)	4.12*	74.8	141.6	72.9
	(22.1)	(411.9)	(534.0)	(401.9)
	[97]	]117]	[72]	[61]

Notes: Point estimates are mean. Standard deviations are in parentheses and the numbers of observations are in brackets. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. The numbers of observations differ across crops because the number of crops grown differs by household. 1 USD = 65.1 INR (Indian rupees).

Although there is an increase in the value of production in the midline survey, the value of production does not differ significantly between treatment and control households. Plot-level

results show that the average yield of rice increased from 1,500 kg/ha to 2,500 kg/ha in the midline survey, maize yield increased from 1,500 kg/ha to 2,500 kg/ha, but Mandia (finger millet) yield was reduced by one-half in the midline survey. We show statistics of these three crops as they are the most widely grown crops by sample households, with rice being the most widely grown crop and Mandia the least.

Panel B of Table 17 shows that, on average, sample households sold output worth around 3,600 rupees at baseline and 11,000 rupees in the midline survey during the Kharif seasons. Compared to the value of production, only 8% of the farm output or value of production was sold at baseline, while around 17% of the farm output was sold in the midline survey, which suggests that sample households mainly produce for home-consumption and have low market participation. This is also supported by the plot-level results which show very low sales values per hectare, especially for rice. More maize is sold per ha compared to rice, even though more rice is sold in absolute terms, probably because the few households that grow maize sell maize, but consume rice. Given that very few households grow and sell maize, these results are only suggestive due to the small sample size. Comparison of treatment and control households shows similarities in sales, except that at baseline smaller quantities of Mandia were sold by treatment households compared to control households, while in the midline sample, treatment households sell significantly more maize per hectare compared to control households.

In Table 18 we compare the summary statistics of the value of production and sales of PVTG and non-PVTG households. Panel A shows that in both survey rounds, PVTG households have significantly smaller values of production (farm output) compared to non-PVTG households. At baseline, PVTG households have a lower rice yield compared to non-PVTG households. However, in the midline survey round, PVTG and non-PVTG households have similar rice yields, which may imply an improvement in the rice yield among PVTG households in the midline survey. Maize and Mandia yields are similar among PTVG and non-PVTG households at baseline, but in the midline survey round, maize yield is significantly lower among PVTG households compared to non-PVTG households. Panel B of Table 18 also shows that household- and plot-level sales statistics are similar for both PVTG and non-PVTG households at baseline, but in the midline sample, PVTG households have a significantly lower value of sales and kilograms of rice sales per hectare.

Table 18. Agricultural yield and sales in Kharif season by tribal groups

	2017		2021	
	PVTG	Non-PVTG	PVTG	Non-PVTG
Panel A: Produced output				
Household-level				
Value of total production (rupees)	39,441.1***	52,207.5	50,531.5***	70,061.7
	(45,797.6)	(55,909.0)	(68,950.2)	(84,469.7)
	[704]	[1,395]	[636]	[1285]
Plot-level				
Paddy yield (kg/ha)	1,337.7***	1,620.2	2,456.2	2,623.2
	(1,449.0)	(1,391.7)	(1,267.7)	(1,149.7)
	[410]	[991]	[856]	[2,026]
Maize yield (kg/ha)	1,608.3	1,536.0	1,796.4***	2,525.9
	(1,644.5)	(1,382.4)	(1,314.5)	(1,263.6)
	[45]	[108]	[49]	[243]
Mandia (finger millet) yield (kg/ha)	1,063.3	894.1	378.1	505.5
	(1,309.1)	(1,116.5)	(638.6)	(693.0)
	[119]	[95]	[63]	[70]
Panel B: Sold output				
Household-level				
Value of total sales (rupees)	3,386.3	3,780.1	7,457.1***	13,036.6
	(8,784.5)	(8,830.2)	(15,763.1)	(21,381.2)
	[704]	[1,395]	[636]	[1,285]
Plot-level				
Paddy sold (kg/ha)	111.8	155.2	78.1***	145.4
	(505.0)	(623.6)	(462.1)	(604.7)
	[410]	[991]	[856]	[2,026]
Maize sold (kg/ha)	293.8	156.0	1,126.1	1,396.3
	(1,011.8)	(783.4)	(1,470.7)	(1,633.8)
	[45]	[108]	[49]	[243]
Mandia (finger millet) sold (kg/ha)	16.3	75.9	54.9	159.8
	(162.1)	(421.5)	(333.4)	(575.2)
	[119]	[95]	[63]	[70]

Notes: Point estimates are mean. Standard deviations are in parentheses and the numbers of observations are in brackets. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. Number of observations differs across crops because the number of crops grown differs by household. 1 USD = 65.1 INR (Indian rupees).

4.1.10 Livestock ownership

Table 19 presents summary statistics of livestock ownership by treatment status. About 80–85% of the sample (both groups) owned livestock at baseline, while in the midline survey around 73–74% of the sample owned livestock. Sample households own around three Tropical Livestock Units (TLUs).[8] While 20% of households received income from animal or animal-product sales at baseline, this rose to about 31–32% of households (both groups) at midline. These low figures suggest that households keep livestock largely for purposes of own-consumption, although they are also engaged with the market and earning income.

Table 19. Livestock ownership and animal product sales

	2017		2021	
Livestock ownership	Treatment	Control	Treatment	Control
Own livestock in the last 12 months (%)	81.1**	85.1	72.2	74.1
	(39.1)	(35.6)	(44.8)	(43.8)
Tropical livestock Units (TLUs)	3.12	3.33	2.97	3.14
	(4.08)	(4.14)	(5.31)	(5.47)
Received income from animal/product (%)	20.8	20.4	30.8	31.9
	(40.6)	(40.3)	(46.2)	(46.6)
Number of animals				
Number of cattle	2.06	2.10	3.43	3.51
	(4.03)	(4.48)	(7.10)	(7.24)
Number of calves	1.04	1.20	0.20***	0.44
	(3.42)	(2.90)	(1.24)	(2.42)
Number of buffaloes	0.90	0.79	0.21	0.21
	(2.43)	(2.97)	(1.50)	(1.23)
Number of goats/sheep	1.48*	1.21	2.42	2.48
	(3.53)	(3.22)	(4.70)	(5.46)
Number of pigs	0.48***	0.95	0.22	0.20
	(1.49)	(2.88)	(1.28)	(1.34)
Number of mules/horses	0.45**	0.63	0.0	0.0
	(2.03)	(2.00)	(0.0)	(0.0)
Number of poultry†	0.90	0.76	3.69	3.28
	(2.74)	(2.90)	(5.93)	(5.37)
Observations	1,050	1,049	962	959

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. †Number of poultry includes the number of chickens, ducks, and other similar birds.

A breakdown of the types of livestock owned by treatment and control households shows that treatment households own more sheep/goats, but fewer pigs and mules/horses at baseline,

[8] TLU is a widely accepted measure for counting livestock based on weights assigned as follows: 0.7 for cattle and buffaloes; 0.6 for horses, donkeys, and mules; 0.5 for calves; 0.2 for pigs; 0.1 for goats and sheep; and 0.01 for poultry. This implies, for example, that owning 10 goats is equivalent to owning 1 cattle (IFAD, 2017).

although the differences are slight (e.g., both groups own fewer than one pig, mule/horse, or chicken at baseline). There were some changes in livestock ownership by midline, although the changes tended to be in the same direction and of roughly the same magnitude for both treatment and control groups. Ownership of buffalo, pigs, and horses/mules went down, but ownership of goats/sheep and poultry went up. Part of OPELIP's goal is to encourage program beneficiaries to take up goat and poultry rearing to help in diversification of their diets and income sources, hence this could explain the increase in numbers of goats/sheep and poultry.

Table 20 presents summary statistics of livestock ownership differentiated between PVTG and non-PVTG households. In both survey rounds, a significantly smaller share of PVTG households compared to non-PVTG households own livestock and received income from livestock sales.

Table 20. Livestock ownership and sales by tribal groups

	2017		2021	
Livestock ownership	PVTGs	Non-PVTGs	PVTGs	Non-PVTGs
Own livestock in the last 12 months (%)	79.8***	84.8	66.8***	76.3
	(40.2)	(35.9)	(47.1)	(42.5)
Tropical Livestock Units (TLUs)	2.72***	3.47	2.50***	3.33
	(3.43)	(4.39)	(4.85)	(5.62)
Received income from animal/product (%)	17.0***	22.4	24.7***	34.6
	(37.6)	(41.7)	(43.2)	(47.6)
Number of animals				
Number of cattle	1.60***	2.32	2.87**	3.76
	(3.24)	(4.67)	(6.49)	(7.47)
Number of calves	1.17	1.09	0.14***	0.41
	(3.83)	(2.78)	(1.49)	(2.10)
Number of buffaloes	0.66**	0.94	0.27	0.18
	(1.73)	(3.08)	(1.66)	(1.21)
Number of goats/sheep	1.31	1.37	1.75***	2.79
	(3.17)	(3.48)	(3.55)	(5.67)
Number of pigs	0.91***	0.61	0.14*	0.25
	(2.66)	(2.09)	(0.85)	(1.48)
Number of mules/horses	0.38***	0.63	0.0	0.0
	(1.07)	(2.35)	(0.0)	(0.0)
Number of poultry†	0.81	0.85	3.00***	3.72
	(2.24)	(3.07)	(5.08)	(5.92)
Observations	704	1,395	636	1,285

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < .10$, $^{**}p < .05$, and $^{***}p < .01$. †Number of poultry includes the number of chickens, ducks, and other similar birds.

PVTG households also have fewer TLUs, cattle, and pigs in both survey rounds compared to non-PVTG households. PVTG households also owned fewer buffaloes and mules at baseline, and fewer calves, goats/sheep, and poultry in the midline sample. Overall, these results show that PVTG households are worse-off in terms of livestock ownership compared to non-PVTG

households, implying that OPELIP has not improved livestock ownership. As mentioned, part of OPELIP's goal is to improve livestock ownership, especially poultry and goats/sheep among PVTG households (IFAD, 2014). The descriptive results, although suggestive, do not show an improvement in production of goats/sheep and poultry both when comparing samples by tribal group or treatment status.

4.1.11 Well-being indicators

Table 21 presents summary statistics of selected well-being indicators differentiated by treatment status. The well-being indicators include income, consumption expenditures, and asset accumulation measures.

Table 21. Well-being indicators by treatment status

	2017		2021	
Income	Treatment	Control	Treatment	Control
Household income, monthly (rupees)	2,001.8	1,956.9	5,942.5*	5,161.3
	(3,901.1)	(3,979.9)	(10,544.9)	(7,385.9)
Income (rupees)†	503.4	491.6	1,268.4	1,196.7
	(902.4)	(866.9)	(1,202.2)	(1,213.6)
Household has no cash income (%)	38.4	37.3	6.65	7.82
	(48.7)	(48.4)	(24.9)	(26.9)
Income poverty (%)	77.4	76.6	40.6*	44.4
	(41.8)	(42.3)	(49.1)	(49.7)
Consumption expenditures				
Household consumption expenditures per month (rupees)	2,801.0	2,925.7	3,304.6	3,199.1
	(2,578.1)	(2,583.4)	(2,557.5)	(2,729.4)
Consumption expenditures, monthly per adult equivalent	780.0	811.1	897.1	892.2
	(744.1)	(772.7)	(665.0)	(747.0)
Consumption poverty (%)	62.1***	56.1	48.2	48.9
	(48.5)	(49.6)	(50.0)	(50.0)
Asset				
Asset Index (score)	0.016	-0.016	0.026	-0.026
	(1.10)	(0.89)	(1.14)	(0.83)
Asset poor (%)	45.9*	41.8	36.1	35.5
	(49.9)	(49.4)	(48.0)	(47.9)
Observations	1,050	1,049	962	959

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. †Income and consumption expenditures are in rupees per-adult equivalent per month.

Income is computed based on 12-month data on gross income from all farm and off-farm economic activities (self-employment, wage employment, social transfers or gifts, income from the sales of livestock and agricultural products, loans and credit received, and remittances) following the same approach used at baseline.

Results show that at baseline, sample households' monthly and per adult equivalent monthly incomes (per adult equivalents rather than per capita comparisons are used for proper comparison of households) are about 2,000 and 500 rupees, respectively, and do not differ by treatment status. At midline, incomes more than doubled, with treatment households having significantly higher monthly incomes than control households. The share of households without cash income dropped from about 37% at baseline to around 7% in the midline sample. Similarly, the proportion of poor households dropped from around 77% to about 45% in the midline survey, with a significantly smaller share of treatment households falling below the income poverty line compared to control households.[9] These results suggest that treatment households have higher incomes and are less likely to be poor in the midline survey, possibly due to participation in OPELIP.

Table 21 also shows that household and per adult equivalent monthly consumption expenditures were about 2,900 and 800 rupees, respectively, at baseline. A larger share of treatment households compared to control households was 'consumption poor' or spent below the poverty line. In the midline survey, treatment and control households do not differ significantly in terms of their monthly consumption expenditures, consumption expenditures per adult equivalent, or 'consumption poverty'.

Lastly, we measure wellbeing using an asset index computed using principal component analysis. The asset index used is a weighted index of a number of durable assets, livestock, and housing characteristics. Unlike income and consumption which reflect short-term well-being status, the asset index captures the current and long-term wellbeing of households (Liverpool-Tasie & Winter-Nelson, 2011; IFAD, 2017). In both survey rounds, the asset index does not differ significantly between the treatment and control households. However, the share of asset-poor treatment households compared to control households is significantly larger at baseline, but not at midline. Although significant differences are not observed in consumption expenditures or in the asset-index between treatment and control groups in the midline sample, the differences in the share of consumption-poor and asset-poor households at baseline become less stark in the midline. This suggests some improvements in these indicators among treatment households.

In Table 22, we compare the summary statistics of income, consumption expenditures, and asset-based well-being measures between PVTG and non-PVTG households. In the two survey rounds analyzed, household and per adult equivalent incomes are lower among PVTG households compared to non-PVTG households. The shares of PVTG households that are income-poor and those that have no cash income are also higher than those of non-PVTG households. PVTG households are also worse-off in terms of consumption expenditures and asset-based indicators as shown by their significantly lower consumption expenditures, higher consumption poverty prevalence rates, lower asset index, and higher asset-poverty prevalence rates compared to PVTG households. While the results in Table 21 show that there

[9] The income poverty line used is 625 rupees per month per person following the baseline analysis (IFAD, 2017). The threshold was selected as it was the per capita monthly poverty cut-off for households in Odisha.

are some improvements in the incomes of treatment households in the midline survey, it seems that huge differences still exist between PVTG and non-PVTG households in terms of well-being indicators that the program has not been able to bridge. This may suggest special attention still needs to be given to PVTG households to lift them out of poverty and improve their wellbeing.

Table 22. Well-being indicators by tribal groups

	2017		2021	
Income	PVTG	Non-PVTG	PVTG	Non-PVTG
Household income, monthly (rupees)	1,867.1	2,036.0	4,137.0***	6,253.1
	(3,989.6)	(3,914.7)	(8,704.2)	(9,230.4)
Income (rupees)†	458.8	517.1	891.6***	1,401.4
	(904.6)	(874.0)	(952.0)	(1,283.8)
Household has no cash income (%)	40.3*	36.6	10.5***	5.60
	(49.1)	(48.2)	(30.7)	(23.0)
Income poor (%)	80.0**	75.6	54.7***	36.5
	(40.0)	(43.0)	(49.8)	(48.2)
Consumption				
Consumption expenditures, per month (rupees)	2,715.9*	2,937.7	2,878.1***	3437.0
	(2,451.9)	(2,641.4)	(2,390.4)	(2,743.8)
Consumption expenditures, monthly per adult equivalent	735.4***	825.9	779.5	951.7
	(671.1)	(797.4)	(662.3)	(721.5)
Consumption poor (%)	63.5***	56.9	56.8***	44.5
	(48.2)	(49.5)	(49.6)	(49.7)
Asset				
Asset Index (score)	-0.074**	0.037	-0.12***	0.062
	(0.90)	(1.04)	(0.78)	(1.09)
Asset poor (%)	45.2	43.7	43.7***	31.8
	(49.8)	(49.6)	(49.6)	(46.6)
Observations	704	1,395	636	1,285

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$, †Income and consumption expenditures are in rupees per-adult equivalent per month.

4.1.12 Income/livelihood diversification

Table 23 presents summary statistics of the number of income sources and income diversity by treatment status and tribal group. Income diversity is measured using the Margalef index

(MI) following FAO (2015)[10]. The index has a lower limit of zero if only one unit of diversity (source of income) is observed and has no limit on the upper bound. A higher MI score implies higher income diversity. Here, the number of income sources and income diversity are used as proxies for household livelihood diversification and resilience. Based on the literature, livelihood diversification has the potential to enhance the resilience of households (Andersen & Cardona, 2013; Wan et al., 2016; Hertel et al., 2021).

Table 23. Income and livelihood diversification by treatment and tribal group

	2017		2021	
Panel A: Treatment status	Treatment	Control	Treatment	Control
Number of income sources (0–6)	0.85	0.86	2.22*	2.11
	(0.89)	(0.91)	(1.20)	(1.19)
	[1,050]	[1,049]	[962]	[959]
Income diversity (Margalef index)	0.05	0.05	0.13	0.12
	(0.06)	(0.07)	(0.09)	(0.09)
	[604]	[605]	[892]	[874]
Panel B: Tribal group	PVTGs	Non-PVTGs	PVTGs	Non-PVTGs
Number of income sources (0–6)	0.81	0.88	1.86***	2.32
	(0.87)	(0.91)	(1.12)	(1.20)
	[704]	[1,395]	[636]	[1,285]
Income diversity (Margalef index)	0.04	0.05	0.10***	0.13
	(0.06)	(0.07)	(0.08)	(0.09)
	[395]	[814]	[566]	[1,200]

Notes: Point estimates are sample means. Standard deviations are in the parentheses. Numbers of observations are in square brackets. A two-sample t-test is used for the test of mean differences between treatment/PVTG and control/non-PVTG groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Panel A of Table 23 shows that the sample households in both groups at baseline have about one income source and that their income diversity score or Margarlef index is low, about 0.05 on average, for both treatment and control households. The number of income sources and income diversity increase in the midline survey to about 2 and 0.13, respectively. At baseline, sample households do not differ by treatment or tribal group in terms of the two indicators. However, in the midline sample, treatment households have slightly more income sources compared to the control households. PVTG households also have fewer income sources and lower income diversity compared to non-PVTG households.

[10] $MI_i = \frac{S_i - 1}{\ln(N_i)}$, where S_i is the number of farmer-managed units of diversity for household i (total number of household income sources) and N_i is the total population count over all farmer-managed units of diversity options (total household income) (FAO, 2015).

4.1.13 Food security outcomes

Table 24 presents results of the food security status of households by treatment status. Food security is measured using an adaptation of the Food Insecurity Experience Scale (FIES). The FIES is a comprehensive tool to understand food security and hunger status over the previous 12 months at the individual or household level, but it does not reflect the food security status of children (Ballard, Kepple, & Cafiero, 2013; IFAD, 2017). The FIES is based on eight questions. Hence, the resulting score ranges from 0 to 8. Households or individuals with a score of 0 are considered food secure. A FIES score between 1 and 2 indicates mild food insecurity, 3 to 5 indicates moderate food insecurity, and 6 to 8 indicates severe food insecurity (Ballard, Kepple, & Cafiero, 2013; IFAD, 2017)[11].

Table 24. Food security by treatment status

Food security status	2017		2021	
	Treatment	Control	Treatment	Control
Panel A: Food insecurity experience scale (FIES)				
FIES score (0–8)	0.9	0.8	1.1**	0.9
	(1.7)	(1.5)	(2.1)	(1.8)
No food insecurity (%)	69.1	66.9	65.0	67.5
	(46.2)	(47.1)	(47.7)	(46.9)
Mild food insecurity (%)	19.3*	22.3	18.6	18.5
	(39.5)	(41.7)	(38.9)	(38.8)
Moderate food insecurity (%)	7.4	8.0	8.9	8.7
	(26.2)	(27.2)	(28.5)	(28.1)
Severe food insecurity (%)	4.1*	2.8	7.5*	5.4
	(19.8)	(16.4)	(26.3)	(22.7)
Panel B: Household Dietary Diversity				
HDDS (0-12) (24 hr)[1]	5.3	5.3	5.8**	6.0
	(1.3)	(1.3)	(1.2)	(1.1)
Observations	1,050	1,049	962	959

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. [1]Food groups: cereals, white roots and tubers, vegetables, fruits, meat, eggs, fish and seafood, legumes, nuts and seeds, milk and milk products, oil and fats, sweets and condiments. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Panel A of Table 24 shows that households have an average FIES score between 0.8 and 1.1 in the two survey periods, which implies that on average the sample households experience mild food insecurity. Although treatment households have a significantly larger FIES score compared to control households in the midline survey, both groups experience mild food insecurity, on average. A breakdown of the FIES score shows that around 67% of the household members in both survey rounds are food secure, around 20% experience mild food

[11] For more on the FIES, see FAO's FIES website: https://www.fao.org/in-action/voices-of-the-hungry/fies/en/.

insecurity, and less than 10% experience moderate or severe food insecurity. Compared to the control group, a smaller share of household members in the treatment group experienced mild food insecurity, while a larger share of the treatment group experienced severe food insecurity at baseline. In the midline survey, a significantly larger share of treatment households also experienced severe food insecurity.

Panel B of Table 24 shows summary statistics of the Household Dietary Diversity Score (HDDS) by treatment status. HDDS is measured as a simple count of 12 food groups consumed by the household 24 hours before the survey following the baseline procedure and FAO guidelines (Kennedy et al. 2010). HDDS is a robust indicator of household food security due to its positive correlation with household calorie/energy intake (Leroy et al. 2015)[12]. On average, sample households consumed close to an average of five food groups at baseline and almost six at midline. Consumption at baseline is almost the same, while treatment households consumed significantly fewer food groups compared to control households at midline, though the difference was very small (0.2 food groups).

Table 25 presents a comparison of the food security status of households by the tribal group. PVTG and non-PVTG households have similar FIES scores in both survey rounds. The shares of PVTG and non-PVTG households that experience zero food insecurity, mild, moderate, or severe food security do not show a clear pattern for all the indicators. For instance, the share of PVTG households that experience zero food insecurity is significantly lower at baseline, but higher in the midline survey round. The scenario is reversed in the case of moderate food insecurity, where a larger share of PVTG households experience moderate food insecurity at baseline, but in the midline survey, the share is smaller compared to that of non-PVTG households. Households consume around five food groups, on average at baseline, while around six at the midline. PVTG households consume significantly lower numbers of food groups in the midline survey compared to Non-PVTGs.

[12] The results differ compared to the baseline report as the food item allocation was corrected within the current analysis. Both baseline and midline calculations are based on the same food item allocation to the twelve respective food groups defined by the guidelines.

Table 25. Food security status by tribal groups

Food security status	2017		2021	
	PVTGs	Non-PVTGs	PVTGs	Non-PVTGs
Panel A: Food insecurity experience scale (FIES)				
FIES score	0.9	0.8	1.1	1.0
	(1.6)	(1.6)	(2.2)	(1.8)
No food insecurity	64.1***	70.0	69.3**	64.7
	(48.0)	(45.8)	(46.1)	(47.8)
Mild food insecurity	22.9	19.8	16.2*	19.7
	(42.0)	(39.9)	(36.9)	(39.8)
Moderate food insecurity	9.5**	6.8	5.7***	10.4
	(29.4)	(25.2)	(23.1)	(30.5)
Severe food insecurity	3.6	3.4	8.8***	5.3
	(18.5)	(18.1)	(28.4)	(22.4)
Panel B: Household Dietary Diversity				
HDDS (0-12) (24 hr)[1]	5.3	5.3	5.6***	6.0
	(1.2)	(1.3)	(1.2)	(1.078)
Observations	704	1,395	636	1,285

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG in the same year. [1]Food groups: cereals, white roots and tubers, vegetables, fruits, meat, eggs, fish and seafood, legumes, nuts and seeds, milk and milk products, oil and fats, sweets and condiments. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.1.14 Dietary quality

Table 26 presents summary statistics of minimum dietary diversity for children (MDD), MDD for women (MDD-W), and MDD for men (MDD-M). MDD and MDD-W are often used as proxies for micronutrient adequacy of children's and women's diets respectively (WHO &UNICEF, 2021; FAO & FHI 360, 2016; FAO, 2021). MDD for children is defined as the share of children 6–23 months who consumed at least five out of eight food groups in the 24 hours before the survey (WHO & UNICEF, 2021). MDD-W is an indicator of whether women 15 to 49 years of age consumed at least five out of ten selected food groups in the 24 hours before the survey (FAO, 2021). MDD-M can be defined in the same fashion as MDD-W, but the questions are asked of men because we wanted to compare men and women's diets using the same indicator. This is a novel contribution which has not been explored in the literature. We do not have any a priori reason to believe that applying the MDD-W questionnaire to men would produce significant inaccuracies (even though we acknowledge that micronutrient needs are different), but further research on this is warranted. All of these indicators were collected only in the midline survey. We additionally, present the MDD-W twice, once for all women within the sample and specifically for women within the validated target range of the MDD-W (women of reproductive age; 15-49 years), as MDD-W WRA. In addition, Table 26 presents the food group score for each indicator, displaying the average number of food groups consumed by children, women and men.

Table 26. Child and adult dietary quality indicators, 2021

	Full sample	Treatment	Control	PVTG	Non-PVTG
	Mean/ (SD)/[Count]	Mean/ (SD)/[Count]	Mean/ (SD)/[Count]	Mean/ (SD)/[Count]	Mean/ (SD)/[Count]
MDD - Ch (%)	35.5	32.5	38.5	42.3	32.0
	(48.0)	(47.1)	(49.0)	(49.9)	(46.9)
	[155]	[77]	[78]	[52]	[103]
MDD-Ch score (0-8) [1]	3.9	3.7	4.1	4.1	3.8
	(1.6)	(1.6)	(1.6)	(1.8)	(1.5)
	[155]	[77]	[78]	[52]	[103]
MDD-W all (%)	19.9	19.4	20.4	22.6**	18.5
	(39.9)	(39.6)	(40.3)	(42.0)	(38.8)
	[1,840]	[918]	[922]	[604]	[1,236]
MDD-W (score) for all women (0-10) [2]	3.3	3.3	3.3	3.1***	3.4
	(1.5)	(1.5)	(1.5)	(1.6)	(1.4)
	[1,840]	[918]	[922]	[604]	[1,236]
MDD-W WRA (%)	19.9	20.1	19.8	24.2***	17.6
	(40.0)	(40.1)	(39.9)	(42.9)	(38.1)
	[1,365]	[692]	[673}	[483]	[882]
MDD-W (score) for women 15-49 years (0-10) [2]	3.3	3.3	3.3	3.2**	3.4
	(1.5)	(1.5)	(1.5)	(1.7)	(1.4)
	[1,365]	[692]	[673}	[483]	[882]
MDD-M (%)	20.8	20.5	21.0	24.7***	18.9
	(40.6)	(40.4)	(40.8)	(43.2)	(39.2)
	[1,678]	[836]	[842]	[546]	[1,132]
MDD-W (score) for men (0-10) [2]	3.3	3.3	3.3	3.2***	3.4
	(1.5)	(1.5)	(1.5)	(1.7)	(1.4)
	[1,678]	[836]	[842]	[546]	[1,132]

Notes: Point estimates are sample means. Standard deviations (SD) are in parentheses. Number of observations (N) in square brackets. MDD-Ch; minimum dietary diversity for children, is defined as % of children 6–23 months who consumed at least five out of eight food groups in the 24 hours before the survey. MDD-W WRA as % of women 15 to 49 years of age who consumed at least five out of ten defined food groups in the 24 hours before the survey. MDD-M is defined in the same manner as MDD-W but focuses on men. [1]Food groups: Breastmilk, grains, roots, tubers and plantains, pulses, dairy products, flesh foods, eggs, vitamin a rich fruits and vegetables and other fruits and vegetables. [2]MDD-W food groups (10): grain, white roots and tubers, plantain, pulses, nuts and seeds, dairy, meat, poultry and fish, egg, dark green leafy vegetables, vitamin A rich fruits and vegetables, other vegetables and other fruits. Level of significance *p < 0.10, **p < 0.05, and ***p < 0.01.

On average, about 36% of children in the sample consume at least five foods groups out of eight possible food groups (fewer in the treatment than the control group, at 33% versus 39%, respectively), while around 21% and 20% of men and women in both groups consume at least five out of ten selected food groups. A slightly higher percentage of men compared to women

reached minimum consumption diversity, but the figures are very close and indicate that a large proportion of adults – around 80% – are not consuming an adequately diverse diet.

In terms of food groups, on average children consume 3.91 different food groups, the control group slightly more compared to the treatment group. This pattern holds if looked into the food group score of women (all and within the age group of reproductive age 15-49 years) and men. On average, all women and women within reproductive age consumed 3.31 different food groups in the past 24 hours, while men consumed 3.33 different food groups. Differences are seen between the treatment and control groups, but none are statistically significant, which differs when we investigate adult food group consumption and tribal groups. Men and women within the non-PVTG group consume statistically significantly more food groups compared to the PVTG group.

The MDD result for our sample (36%) is slightly higher than the recent MDD figure of 31% observed for children surveyed between 2016 and 2018 in the state of Odisha (CNNS, 2019). While it still means that a large majority of children are not eating adequately, it also indicates that children are in some way having more adequately diverse diets than adults. Therefore, these results suggest that the majority of children, women, and men in the sample do not consume a diet with adequate micronutrients. MDD-Ch, MDD-W, and MDD-M do not differ significantly by treatment status but for adults differ statistically significantly by the tribal group. Significantly larger shares of women and men in PVTG households consume diets with adequate micronutrients compared to those in non-PVTG households.

4.1.15 Anthropometric indicators

Table 27 presents summary statistics of anthropometric indicators – prevalence of stunting, underweight, wasting and overweight – for children between 0 and 59 months of age. We compare households by treatment status and tribal group. The calculations used for the current analysis differ compared to the ones within the baseline report. The results in Table 27 are computed using the WHO standard procedure (inclusive cut-off points) and Stata tool (WHO, 2011; WHO Multicentre Growth Reference Study Group, 2006)[13].

In 2017, 47% of the sampled children were stunted, 46% were underweight, and 32% were wasted on average. The differences do not differ significantly between treatment and control households at baseline. In the midline survey, slightly lower stunting rates are observed. The same holds for the prevalence of being wasted and underweight. Treatment households have a statistically significantly lower prevalence of underweight children compared to control households. In both survey rounds, anthropometric indicators are slightly higher, except for wasting, than those reported by the National Family Health Survey (NFHS) 5, 2019–2021 for the state of Odisha, which shows that between 2019 and 2021, 31% of the children under five years of age were stunted, 29.7% were underweight, and 18.1% were wasted (NFHS 5, 2019-2021). This is probably because our study includes PVTGs and STs, which are more vulnerable to malnutrition challenges as they tend to consume less nutritious diets (Choudhury et al., 2020).

Comparison of the households by tribal group shows that prevalence rates of stunting children are statistically significantly higher in PVTG households than in non-PVTG households at

[13] WHO tool retrieved from https://github.com/unicef-drp/igrowup_update using the following cut-off points (outlier flagging): Weight-for-age z-score < -6 SD or > 5 SD; Length/height-for-age z-score < -6 SD or > 6 SD; Weight-for-length/height z-score < -5 SD or > 5 SD

baseline. However, in the midline sample, no statistically significant differences are observed. This may suggest an improvement in the prevalence of underweight children in PVTG households given that they were worse off than non-PVTG households at baseline. However, this may not be the case since the prevalence of underweight children has not reduced in the midline sample. In addition, comparisons of mild, standard and severe prevalence for stunting, wasting, underweight and overweight of children between 0 and 59 months of age are presented in Table A2.

Table 27. Child anthropometric indicators by treatment status and tribal group

	2017		2021		2017		2021	
	Treatment	Control	Treatment	Control	PVTGs	Non-PVTGs	PVTGs	Non-PVTGs
Children (0-59 months)								
Stunted (<-2 SD)	0.45	0.48	0.43	0.47	0.54*	0.43	0.48	0.43
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Underweight (<-2 SD)	0.43	0.50	0.36**	0.44	0.46	0.46	0.43	0.38
	(0.49)	(0.50)	(0.48)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)
Wasted (<-2 SD)	0.31	0.33	0.14	0.16	0.37	0.29	0.18	0.13
	(0.46)	(0.47)	(0.35)	(0.37)	(0.48)	(0.46)	(0.38)	(0.34)
Overweight (>+ 2 SD)	0.06	0.05	0.01**	0.04	0.04	0.06	0.02	0.02
	(0.23)	(0.21)	(0.10)	(0.19)	(0.20)	(0.23)	(0.15)	(0.15)
Observations	180	191	288	259	125	246	208	339

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between treatment/PVTG and control/non-PVTG in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 28 presents summary statistics of anthropometric indicators for adults. This analysis is based on the midline sample since these indicators were not collected at baseline. Body mass index (BMI) is computed as weight in kilograms divided by the height in meters squared. Adults are classified as underweight if their BMI is less than 18.5. For normal weight, the BMI range is 18.5 – 24.9, while overweight adults have a BMI of more than 24.9 (WHO, 2022). In both samples, the BMI is around 20, which implies that, on average, both men and women in the sample are of normal weight. In both samples, at least 20% of men and women are underweight; the larger share (at least 60%) of men or women is of normal weight, and less than 10% of men and women are overweight. Among men, the indicators are similar between the treatment and control groups. While the BMI score is almost similar in both groups, a statistically significantly larger share of women in the treatment group is underweight, and the share of women of normal weight is statistically significantly smaller compared to the control group.

Table 28. Adult anthropometric indicators by treatment status in 2021

	Men		Women	
	Treatment	Control	Treatment	Control
Body mass index (score)	20.7	20.6	20.1	20.2
	(2.7)	(2.7)	(3.1)	(2.9)
Prevalence of underweight (%)	21.6	22.6	31.8***	27.2
	(41.2)	(41.8)	(46.6)	(44.5)
Prevalence of normal weight (%)	71.1	71.7	61.6**	66.5
	(45.4)	(45.1)	(48.7)	(47.2)
Prevalence of overweight (%)	6.8	5.5	6.3	6.1
	(25.2)	(22.9)	(24.2)	(23.9)
Observations	823	833	912	919

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between treatment and control groups in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$

In Table 29, we compare adult anthropometric indicators by tribal group. In both women and men, PVTG households have significantly lower adult BMIs and larger shares of underweight adults compared to non-PVTG households. A significantly smaller share of PVTG men is of normal weight compared to non-PVTG men. The prevalence of overweight is similar among men in PVTG and non-PVTG households, but the prevalence of overweight women is lower in PVTG households compared to non-PVTG households. Overall, these results suggest that women and men in PVTG households are marginally worse-off in terms of anthropometric indicators.

Table 29. Adult anthropometric indicators by tribal group in 2021

	Men		Women	
	PVTG	Non-PVTG	PVTG	Non-PVTG
Body mass index (score)	20.4**	20.7	19.8***	20.3
	(2.6)	(2.7)	(3.0)	(3.0)
Prevalence of underweight (%)	25.1**	20.7	31.6	28.5
	(43.4)	(40.5)	(46.5)	(45.1)
Prevalence of normal weight (%)	69.4	72.3	63.2	64.5
	(46.1)	(44.8)	(48.3)	(47.9)
Prevalence of overweight (%)	5.1	6.7	4.8	6.8
	(22.0)	(25.0)	(21.5)	(25.2)
Observations	530	1126	598	1233

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences between PVTG and non-PVTG samples in the same year. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$

4.1.16 Village-level infrastructure

The community empowerment and community infrastructure and drudgery reduction components aimed to: promote village development associations (VDAs) to plan need-based activities; build the capacity of VDAs by providing training to plan and execute the needs of

the community, which include natural resource management, community-based paralegal services, healthcare, hygiene and nutrition education, and community infrastructure; and upgrade village link roads, SHG work sheds, toilets, and gravity-fed clean drinking water supply facilities. Against this background, we also assess village-level characteristics (villages' access to markets, irrigation, roads, and public transportation) by treatment status as shown in Table 30.

Each village is considered as a single unit and the center of the village is the reference point for all distances. Village-level questions were asked to presiding village leaders using a village-level questionnaire. In Panel A, the baseline results show that about 12% of the villages in both treatment and control groups had an agricultural output market, which was located about 10 km away. However, this declined substantially by midline with only 4% of treatment villages and 7% of control villages having an output market where households could sell agricultural products about 8 to 9 kilometers away.

Villages in treatment and control groups at baseline had an input market (12% and 18%, respectively), which was located about 8-9 kilometers away. Although the situation stayed about the same for the treatment group at midline, the percentage of the control group having an input market declined to 11%. About 17% of treatment villages had a fair price shop at baseline, as did 24% of control villages. Comparison of treatment and control households shows that treated households were located somewhat further from the nearest fair price shop (4 versus 3 kilometers). However, access to fair price shops fell substantially between baseline and midline, to 10% and 9%, respectively, for treatment and control villages. And the distance for both groups increased to 6 kilometers.

Panel B shows results of access to irrigation, measured as the share of irrigated land in wet and dry seasons. Around 60% of land was irrigated for both groups at baseline during the wet season. This figure rose to 83–84% at midline. In the dry season, however, about 10% of land for both groups was irrigated at baseline, which rose to 23–26% at midline. It is not clear what the reasons are for the differences. Perhaps limited availability of water during the dry season discourages irrigation as long distances have to be travelled to obtain water.

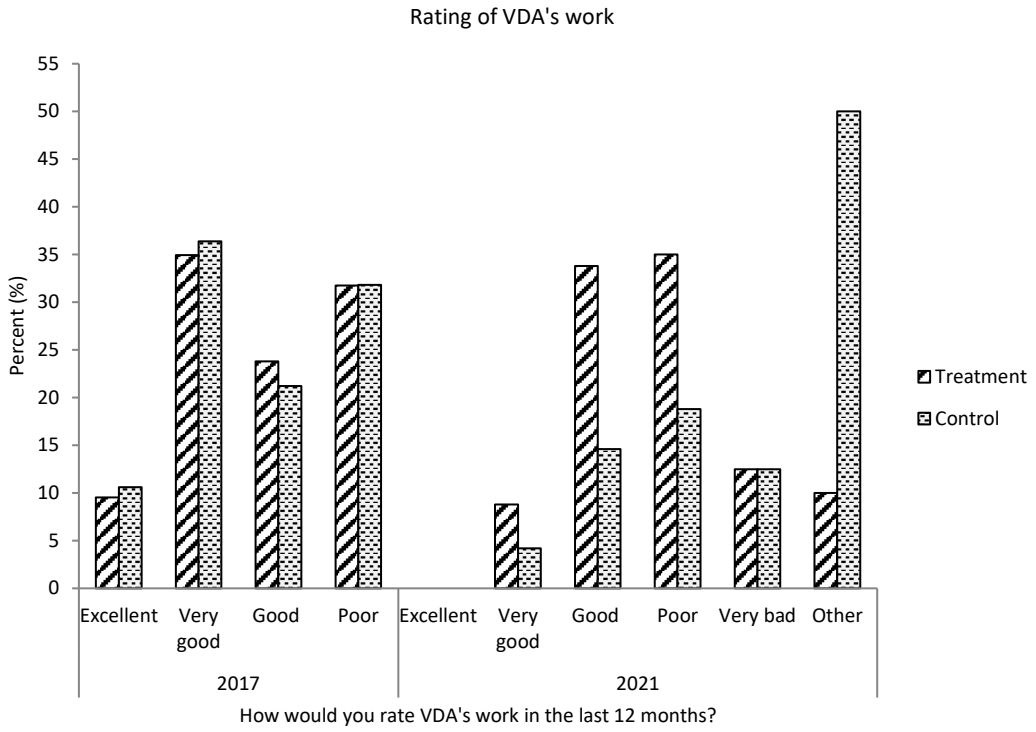
In panel C, we examine access to paved roads and public transport. Less than 10% of the villages had access to public transport at baseline, while 32% of treatment and 38% of control villages did a midline survey, which suggests an improvement in access to public transport. Treated households are located somewhat further away than control households from public transportation at baseline (about 1.8 kilometers), but not in the midline survey (there remains a difference, but it is no longer significant). On average, public transport is located about 4–6 kilometers away and paved roads are located about 2–4 kilometers away in both survey rounds for both groups.

Since strengthening VDA's capacity and service delivery was one of the goals of OPELIP, we also assess VDAs work and the respondents' (village leaders') view on the service provided by VDAs. Panel D shows that, on average, leaders from 47% of the villages in both groups believed people were satisfied with the VDA's work at baseline. In the midline survey, 39% of the treatment villages were satisfied with VDA's work, while only 9% of the control villages were satisfied. Village leaders also rated VDA's work on a scale of one to four, with one being excellent and four being poor, at baseline. Figure 3 shows that about 10% of the respondents in both treatment and control villages rated VDA's work as excellent, 36% as very good, 21% as good, and 33% as poor. In the midline survey, the ratings were a bit different and had an additional rating of "very bad" and "other" (for those who couldn't rate

Table 30. Access to markets, irrigation, transportation, and VDA's services

	2017		2021	
Panel A: Access to market	Treatment	Control	Treatment	Control
Village has a market to sell agricultural produce (%)	13.5	11.0	3.45	7.41
	(34.3)	(31.4)	(18.4)	(26.4)
Distance to the nearest output market (km)	9.5	10.7	9.29	7.84
	(11.7)	(15.2)	(7.88)	(6.61)
Village has a market to buy agricultural inputs (%)	12.4	17.6	12.6	11.1
	(33.1)	(38.3)	(33.4)	(31.6)
Distance to the nearest input supplier (km)	9.2	7.9	7.97	7.40
	(6.80)	(6.53)	(14.1)	(6.51)
Village has a fair price shop (%)	16.9	24.2	10.3	8.64
	(37.6)	(43.1)	(30.6)	(28.3)
Distance to the nearest fair price shop	4.1**	2.9	5.62	5.82
	(3.24)	(2.70)	(12.4)	(15.9)
Distance to the nearest wholesale market (km)	14.0	13.8	13.5	12.8
	(11.8)	(9.95)	(12.0)	(11.4)
Panel B: Access to irrigation				
Land irrigated in wet season (%)	60.8	63.8	84.0	83.0
	(37.1)	(35.5)	(13.4)	(11.1)
Land irrigated in dry season (%)	10.1	9.5	23.2	25.6
	(20.6)	(23.9)	(18.0)	(16.8)
Panel C: Access to public transportation and roads				
Public transportation passes by the village (%)	7.9	9.9	32.2	38.3
	(27.1)	(30.0)	(47.0)	(48.9)
Distance to the nearest public transportation (km)	6.6**	4.8	5.50	4.24
	(5.49)	(3.69)	(4.74)	(2.84)
Distance to the nearest paved road (km)	4.1	3.9	2.94	1.95
	(4.43)	(4.44)	(5.43)	(2.80)
Panel D: *VDA's services and office*				
Satisfied with VDA's services (%)	47.19	47.25	39.08***	8.64
	(50.20)	(50.20)	(49.08)	28.27
VDAs have physical offices (%)	66.67	69.69	26.83***	8.70
	(47.52)	(46.31)	(44.58)	(28.38)
VDA offices equipped with furniture and stationery (%)	49.20	58.46	48.15***	11.48
	(50.39)	(49.66)	(50.28)	(32.14)
Observations	89	91	87	81

Notes: Point estimates are sample means. Standard deviations are in the parentheses. A two-sample t-test is used for the test of mean differences. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.



VDA's work). None of the respondents rated VDA's work as excellent in the midline survey. In treatment villages, the VDA's work was rated as follows: 9% very good, 34% good, 35% poor, 13% very bad, and 10% other. In control villages, the ratings were as follows: 4% very good, 15% good, 18% poor, 13% very bad, and 50% other, as control villages may not have had VDAs. One-third of the sample at baseline rated the VDA's work as poor, but in the midline sample, a larger share of respondents rated VDA's work as poor or worse, especially among treatment villages where the share is almost 50%.

When we examined whether VDAs have physical offices with minimal required furniture and stationery, baseline results in Panel A of Table 30 showed that 67–70% of VDAs in both groups had a physical office, but in the midline survey, only 27% of treatment and 9% of control group VDA's had a physical office. At baseline, 49% of VDAs in project areas were equipped with furniture and stationery compared to 58% of the VDAs in control areas. However, in the midline survey, 48% of the VDAs in project areas were equipped with furniture and stationery compared to only 11% of VDAs in control areas. In spite of the important role of VDAs, their ratings and physical presence seem to be declining, especially in the control villages.

4.2 Estimation results

So far, we have presented results of the descriptive statistics for sample households differentiated by treatment and PVTG households. Although the results show some statistically significant differences between treatment and control households on various outcome variables, the results cannot be interpreted as having an impact since they do not account for potential confounding factors. In the following subsections, we present estimation results for selected outcome variables after controlling for potential confounding factors using the identification strategy described in section 3. Our aim is to answer the questions highlighted in subsection 2.2 and cover the outcome variables of interest to IFAD namely, the value of production, value of market sales, household income, consumption, asset ownership,

nutrition and food security, land tenure, gender (e.g., women's empowerment), and resilience (livelihood diversification).

4.2.1 Effects of OPELIP on agricultural inputs, value of production, and sales revenue

Table 31 presents estimation results of the effects of OPELIP/treatment on agricultural inputs, value of production, and revenue (sales). Panel A of Table 30 shows that being in the treatment group has a positive and significant effect on the use of chemical fertilizer, but the effects on the use of improved seed, pesticide, and manure are statistically insignificant. In particular, treatment increases the likelihood of using chemical fertilizer by 15.2 percentage points, over-and-above the mean use in the control group. Under the natural resource component, OPELIP aims to increase crop productivity by providing technical training on good agricultural practices, improved soil fertility and soil health management; hence, it is possible that the trainings contributed to increased households' probability of adopting chemical fertilizer (IFAD, 2014). Manure and pesticides are often less used compared to improved seed and fertilizer; hence, the insignificant effects on the inputs are plausible.

Panel B shows the effects of treatment on the value of production (total value, crop, and livestock). Given the large standard deviations of the values of production, we also estimate models with log-transformed values of the dependent variables. In our discussion, we rely on the results of the log-transformed dependent variables as they are less skewed, but also show results of the dependent variables in absolute values. Results show that treatment has positive and significant effects on the (log-transformed) total value of production and value of crop production, but the effects on log-transformed value of livestock production are statistically insignificant. Specifically, the results show that treatment increases the total value of production and value of crop production by 51.1% and 94.5%, respectively.[14] These results are plausible. The increased likelihood of using fertilizer among program beneficiaries shown in Panel A could be contributing to the higher value of crop production, which seems to be driving the overall increase in total production as livestock production does not increase.

In panel C of Table 31, we present estimation results of the effects of OPELIP on the value of sales or revenue (total, crop, and livestock). The sales revenues presented in panel C only represent gross income from agricultural output and not the household's entire gross income.

Also, the values of sales shown can be used as a proxy for the sample households' access to markets or degree of market participation. In this regard, a positive and significant value of sales would suggest a higher degree of market access among treatment households compared to control households and would also imply higher revenues/gross farm income for treatment households. Contrary to our expectations, being in the treatment group does not have significant effects on crops, livestock, or total sales revenue. These results may imply either that although there is an increase in the value of production, the increase is not large enough to contribute to improved sales or that market access is a major challenge among beneficiaries. Indeed, village-level data shows that only 3.5% of the treatment villages have a market for selling agricultural products, which underscores the fact that market access is a major challenge among treatment households.

[14] Percent of the treatment effect for the log-transformed outcome is computed using the formula $100(e^{ATT} - 1)$, where e is the exponential (Kennedy, 1981).

Table 31. Effect of OPELIP on agricultural input use, production, and sales

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Inputs			Use of improved seed (%)	Use of chemical fertilizer (%)	Use of manure (%)	Use of pesticide (%)
Time × Treatment (program)			3.145	15.235***	-2.337	3.690
			(4.013)	(4.235)	(4.802)	(3.555)
Control variables			Yes	Yes	Yes	Yes
R-squared			0.339	0.159	0.319	0.100
Observations			7,028	7,028	7,028	7,028
Panel B: Production	Total value of production (rupees)	Log of total value of production	Value of crop production (rupees)	Log of value of crop production	Value of livestock production (rupees)	Log of value of livestock production
Time × Treatment (program)	3941.3	0.413**	26,908.1*	0.665**	-1,520.8*	0.361
	(2,689.7)	(0.188)	(14,451.2)	(0.262)	(891.8)	(0.242)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.757	0.274	0.020	0.231	0.956	0.288
Observations	4, 020	4, 020	4, 020	4,020	4,020	4,020
Panel C: Revenue/Sales	Value of sales (rupees)	Log of value of sales	Value of crop sales (rupees)	Log of value of crop sales	Value of livestock sales (rupees)	Log of value of livestock sales
Time × Treatment (program)	510.2	0.096	2398.6	0.327	-729.4	-0.180
	(1,471.3)	(0.377)	(1,585.7)	(0.417)	(1,332.5)	(0.289)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.167	0.206	0.218	0.244	0.127	0.168
Observations	4,020	4,020	4,020	4,020	4,020	4,020

Notes: Point estimates are estimated using difference-in-difference regression. Robust standard errors clustered at the village-level appear in parentheses. Control variables include: age of household head, age of head squared, sex of head, marital status of head, education of head, household size, land size, squared land size, land *patta*, irrigation dummy, TLU, group membership, prior coverage by MPA, time, and interaction of time district dummies. 1 USD = 65.1 INR (Indian rupees). Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.2.2 Effects of OPELIP on well-being indicators

Table 32 presents results of the effects of OPELIP on monthly income per adult equivalent, consumption expenditures per adult equivalent, asset accumulation, and poverty. We also show results of econometric models with absolute and log-transformed values of income and expenditure variables due to their large standard deviations.

Columns (2) and (3) of Table 32 show that being in the treatment group increases monthly income per adult equivalent by 50.9% and reduces the probability of 'income poverty' – earning less than 625 rupees per person per month – by 6.5 percentage points, after controlling for confounding factors. Panel B also shows that being in the treatment group has a significant effect on monthly consumption expenditures per adult equivalent and 'consumption poverty' – spending less than 625 rupees per person per month. In particular, treatment increases monthly expenditure per adult equivalent by 11.9% and reduces the probability of being 'consumption poor' by 8.8 percentage points. Interestingly, these results show that OPELIP (being in the treatment group) reduces poverty regardless of whether poverty is measured in terms of income or consumption expenditures, which underlines the robustness of the results. While being in the treatment group has no effect on the asset index, it reduces the probability of being asset poor (falling below the 40^{th} percentile of sample households in terms of asset accumulation) by 7.4 percentage points, after controlling for confounding factors.

Overall, these results show that OPELIP has so far contributed to higher monthly income and consumption expenditures per adult equivalent and has also reduced poverty. These are welcome findings given that the goal of OPELIP is to reduce poverty and improve the wellbeing of PVTGs through income growth and improved food and nutrition security for about 62,356 households, including other tribal and poor households in the intervention areas (IFAD, 2017; IFAD, 2014). A crucial question though is through which mechanism(s) does OPELIP contribute to higher monthly per adult equivalent incomes or consumption expenditures? Given that the effect of OPELIP on gross farm income (sales revenues) is insignificant, it is plausible that the increments are driven by growth in non-agricultural (off-farm) income which accounts for at least 70% of total household income. Indeed, separate analyses (not presented) showed that OPELIP significantly increases off-farm income, but has no effect on farm income. This may mean that greater focus should still be paid to the program's implementation, especially in regard to Component 2, which focuses on natural resource management and livelihoods enhancement to ensure that the program also contributes to increased farm incomes.

Table 32. Well-being indicators

	(1)	(2)	(3)
Panel A: Income	Per adult equivalent monthly income (rupees)	Log of per adult equivalent monthly income	Income poor (%)
Time × Treatment (program)	96.113	0.412*	-6.472**
	(76.160)	(0.214)	(2.883)
Control variables	Yes	Yes	Yes
R-squared	0.278	0.347	0.285
Observations	4,020	4,020	4,020
Panel B: Consumption	Per adult equivalent monthly expenditure (rupees)	Log of per adult equivalent monthly expenditure	Consumption poor (%)
Time × Treatment (program)	81.57	0.113**	-8.812***
	(50.076)	(0.049)	(3.469)
Control variables	Yes	Yes	Yes
R-squared	0.209	0.313	0.214
Observations	4,020	4,020	4,020
Panel B: Asset		Asset index	Asset poor (%)
Time × Treatment (program)		0.079	-7.605*
		(0.068)	(3.100)
Control variables		Yes	Yes
R-squared		0.262	0.218
Observations	4,020	4,020	4,020

Notes: Point estimates are estimated using difference-in-difference regression. Robust standard errors clustered at the village level appear in parentheses. Control variables include: age of household head, age of head squared, sex of head, marital status of head, education of head, household size, land size, squared land size, land *patta*, irrigation dummy, TLU, group membership, prior coverage by MPA, time, and interaction of time district dummies. †Income and consumption expenditures are in rupees per-adult equivalent per month. 1 USD = 65.1 INR (Indian rupees). Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.2.3 Effect of OPELIP on income and livelihood diversification

Table 33 presents the results of the effects of treatment (OPELIP) on the number of income sources and income diversity. As mentioned, the number of income sources and income diversity are used as proxies for livelihood diversification and resilience. Results show that treatment has no effect on the indicators, which implies that there is no evidence that OPELIP is contributing to income/livelihood diversification and resilience. This may also suggest that the program has not contributed to increased income or livelihood opportunities, contrary to what is envisioned in the program's objectives under Component 2.

Table 33. Effect of OPELIP on income and livelihood diversification

	(1)	(2)
Panel A: Income	Number of income sources (0-6)	Income diversification (Margalef index)
Time × Treatment (program)	0.076	0.005
	(0.065)	(0.007)
Control variables	Yes	Yes
Wald Chi or R-squared	2,710	0.337
Observations	4,020	2,975

Notes: Point estimates are estimated using difference-in-difference regression. Robust standard errors clustered at the village level appear in parentheses. Control variables include: age of household head, age of head squared, sex of head, marital status of head, education of head, household size, land size, squared land size, land *patta*, irrigation dummy, TLU, group membership, prior coverage by MPA, time, and interaction of time district dummies. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.2.4 Effect of OPELIP on food security indicators

Table 34 shows the estimation results of the effects of OPELIP (treatment) on food security, measured in terms of FIES, HDDS, and food expenditures. Results show that treatment has no effect on FIES and HDDS, but significantly increases food expenditures per adult equivalent in absolute and log-transformed value terms (columns 3 and 4, respectively). As expected, column 4 shows that being in the treated group increases monthly food expenditures per adult equivalent by 12.1% compared to being in the control group. This is consistent with the higher per adult equivalent incomes and overall consumption expenditures observed among treatment groups. However, the insignificant effect of treatment (OPELIP) on FIES and HDDS is contrary to our expectation that OPELIP would reduce FIES and increase HDDS. One possible reason for the insignificant results could be the low farm production diversity of farm households observed (on average, households grow just a single crop per season as shown in Table 12), which potentially makes it difficult for farm households to diversify their diets through consumption of own-produced foods. Limited access to markets may also prevent program households from diversifying their diets through consumption of purchased nutritious foods. Against this background, it would be important to strengthen access to markets by constructing new markets or rehabilitating existing markets as envisioned in the program's blueprint (IFAD, 2020b). Improved access to the market can spur adoption of improved technologies, boost production, and increase incomes. Promotion of kitchen gardens could also contribute to higher farm and dietary diversity (Pandey et al., 2016).

Table 34. Food security indicators

	(1)	(2)	(3)	(4)
	Food Insecurity Experience Scale (0-8)	Household Dietary Diversity Score (0-12)	Food expenditure per adult equivalent (rupees)	Log of food Expenditure per adult equivalent
Time × Treatment (program)	0.161	0.015	70.11**	0.114**
	(0.140)	(0.021)	(30.03)	(0.048)
Control variables	Yes	Yes	Yes	Yes
(Wald chi2) R-squared	996.6	499.5	0.218	0.351
Observations	4,020	4,020	4,020	4,020

Notes: Point estimates are estimated using difference-in-difference regression. Robust standard errors clustered at the village level appear in parentheses. Poisson regressions are estimated for count variables. Control variables include: age of household head, age of head squared, sex of head, marital status of the head, education of head, household size, land size, squared land size, land *patta*, irrigation dummy, TLU, group membership, prior coverage by MPA, time, and interaction of time district dummies. †Income and consumption expenditures are in rupees per-adult equivalent per month. 1 USD = 65.1 INR (Indian rupees). Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.2.5 Effect of OPELIP on child anthropometric indicators

Table 35 presents the results of the effects of treatment (OPELIP) on child anthropometric indicators, measured in terms of the prevalence of stunting, underweight, and wasting for children under five years of age. Results show that treatment has no effect on any of the three child anthropometric indicators, which implies that there is no evidence to suggest that OPELIP is contributing to improved child nutrition. Estimation results with anthropometric indicators adjusted for outliers also do not show significant effects (Table A3 in the appendix). Together with the results in Table 34 which show insignificant effects of treatment on FIES and HDDS, we conclude that there is no evidence to suggest that treatment (OPELIP) has improved household food security and child nutrition.

Table 35. Effect of OPELIP on child anthropometric indicators

	(1)	(2)	(3)
	Stunting (%)	Underweight (%)	Wasting (%)
Time × Treatment (program)	-3.224	-1.691	-5.392
	(6.121)	(5.683)	(4.926)
Control variables	Yes	Yes	Yes
R-squared	0.126	0.096	0.057
Observations	1,113	1,113	1,113

Notes: Point estimates are estimated using difference-in-difference regression. Robust standard errors clustered at the village level appear in parentheses. Control variables include age of the household head, age of head squared, sex of the head, marital status of the head, education of head, household size, land size, squared land size, land *patta*, irrigation, TLU, group membership, prior coverage by MPA, time, the interaction of time district dummies, quality of drinking water, improved toilet dummy, and quality of energy. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.2.6 Effect of OPELIP on dietary quality and adult anthropometric indicators

Apart from examining the effects of treatment (OPELIP) on child anthropometry, we also evaluate the effects of OPELIP on dietary quality (measured in terms of minimum dietary diversity (MDD) for children aged 6–23 months, MDD for men (MDD-M), MDD for women (MDD-W)), and adult anthropometrics (measured in terms of BMI, the prevalence of underweight, and the prevalence of overweight). For these analyses, we rely on cross-sectional data since data for these indicators were only collected in the midline survey. Table 36 shows that treatment has no effect on MDD for children's or adults' dietary quality and anthropometric indicators. Based on these results, we conclude that there is no evidence to suggest that OPELIP has contributed to improved child or adult nutrition.

Table 36. Effect of OPELIP on dietary quality and adult anthropometrics

	(1)	(2)	(3)
	Neighbor matching	Kernel matching	Observations
	ATT(SE)	ATT(SE)	
MDD (%) [a]	-21.875	-16.227	160
	(29.408)	(30.246)	
MDD-M (%) [b]	6.533	5.628	1,715
	(4.509)	(4.894)	
MDD-W (%) [c]	7.157	6.146	1,855
	(4.624)	(4.067)	
Anthropometrics for men			
Body mass index	-1.058	-1.043	1,701
	(0.790)	(0.764)	
Prevalence of underweight (%)	2.728	1.491	1,701
	(7.096)	(6.661)	
Prevalence of overweight (%)	-2.135	-2.994	1,701
	(6.381)	(6.108)	
Anthropometrics for women			
Body mass index	-0.599	-0.385	1,854
	(1.005)	(0.877)	
Prevalence of underweight (%)	-5.427	-6.089	1,854
	(9.332)	(8.492)	
Prevalence of overweight (%)	-6.076	-5.567	1,854
	(7.453)	(6.609)	

Notes: Point estimates are estimated using propensity score matching. Boot standard errors using 200 replications appear in parentheses. [a] MDD; minimum dietary diversity, [b] MDD-M; minimum dietary diversity for men, [c] MDD-W; minimum dietary diversity for women. Control variables include age of the household head, age of head squared, sex of head, marital status of head, education of head, household size, land size, squared land size, land *patta*, irrigation dummy, TLU, group membership, prior coverage by MPA, district dummies. Anthropometric estimation models include additional variables namely, quality of drinking water, improved toilet dummy, and quality of energy. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.2.7 Effect of OPELIP on other selected indicators

Table 37 shows the estimation results of the effects of treatment on selected outcome variables that include land tenure (having a land title certificate), agricultural practices (slash and burn contour farming, and regular/settled agriculture), child school enrolment (measured in terms of the share (%) of all school-aged children currently attending school), and women's empowerment (share of all decisions made by women alone).

The results show that being in the treatment group significantly increases the probability of having a plot/land title certificate by 5.9 percentage points. Under the natural resource management component, OPELIP aims to improve land tenure by issuing landless households with land title certificates for the land they have been cultivating, therefore the observed results are plausible. The results suggest that the land title certificates issued by the program are contributing to improved ownership of land title certificates (land tenure). At the time of writing this report, more land titles had just been issued but had not been captured in the midline survey, so probably higher improvements will be reported in the endline survey. In terms of agricultural practices, the results show that being in the treatment group does not have an effect on contour farming. However, it reduces the probability of practicing slash and burn by 4.0 percentage points and increases the probability of practicing regular or settled agriculture by 4.7 percentage points. These results are plausible and suggest that the natural resource component of the program is contributing to better agricultural practices among beneficiaries, by reducing slash and burn, and shifting cultivation agriculture (promoting regular agriculture). Improved land tenure could also be contributing to the probability of practicing regular or settled agriculture.

Results also show that being in the treatment group does not have a significant effect on women's empowerment. This is contrary to OPELIP's aim of improving women's empowerment and gender balance through its community empowerment component. Available administrative data shows that as of August 2020, the participation rate in women's empowerment activities was only 28% (IFAD, 2020b). With such low outreach and uptake, it is not surprising that there is no effect of women's empowerment; therefore, it would be important to encourage women to increase their participation in women's empowerment activities for impact to be realized in the endline evaluation.

Analysis of child school enrolment shows that the program has neither stimulated a significant increase in enrollment of children in schools nor reduced child school enrolment. Child school enrolment is analyzed both as a possible unintended outcome that may be affected through increased demand for labor for agricultural production, which may undesirably force children out of school to meet the high demand for agricultural labor, or improved welfare gains that may increase children's enrolment in school. The analysis was also done because, under Component 3, OPELIP focuses on improving community infrastructure, reducing drudgery, and building critical social infrastructure, such as schools, which may possibly increase child school enrolment (IFAD, 2020b).

Table 37. Effect of OPELIP on land tenure, agricultural practices, school enrolment, and women's empowerment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Land title certificate (%)	Contour farming (%)	Slash and burn (%)	Settled or regular agriculture (%)	Women's empowerment (% of women-only decisions) (0-1)	Women's empowerment (% of joint decisions) (0-1)	School-aged children attending school (%)
Time × Treatment (program)	5.906*	-4.032	-4.026*	4.743**	-0.006	0.001	1.155
	(3.267)	(2.641)	(2.212)	(2.295)	(0.007)	(0.018)	(3.187)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.427	0.181	0.152	0.189	0.887	0.294	0.112
Observations	7,754	7,754	7,754	7,754	4,020	4,020	2,739

Notes: Point estimates are estimated using difference-in-differences regressions. Robust standard errors clustered at the village level in parentheses. Control variables include the age of the household head, age of head squared, sex of head, marital status of head, education of head, household size, land size, squared land size, land *patta*, irrigation, TLU, group membership, prior coverage by MPA, and district dummies. Level of significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

5. Conclusions and recommendations

While there has been a huge investment in agricultural development interventions, few rigorous evaluations of such interventions exist and little is known about their impact. In this report, we conducted a midline evaluation of the effects of OPELIP – a program which aims to reduce poverty and improve the wellbeing of PVTGs through income growth and improved food and nutrition security for about 62,356 households, which includes other tribal and poor households in the intervention areas – on various indicators. The midline assessment aims to provide useful mid-program implementation insights and feedback for continued program implementation and design of new aspects for the second half of the program, from 2021 to 2024, as well as for future programs. The assessment also aims to contribute towards IFAD's efforts in generating a large body of evidence on the impacts of rural development interventions on economic development and living standards of poor and vulnerable populations, as well as filling the knowledge gap on the impacts of development programs.

Results from a sample of about 2,000 households analyzed using difference-in-differences and propensity score matching reveal that the program has improved the adoption of fertilizer by 15.2 percentage points, but that it has not had an effect on the adoption and use of other soil fertility and yield-enhancing inputs among program beneficiaries, contrary to the program's objectives. In terms of production, sales, or market access, the results show that the program increased overall agricultural production by 51.1%, with the increase mainly driven by the rise in crop production, as livestock production did not increase. However, OPELIP did not have an effect on agricultural sales revenues, which may imply that the gains in total agricultural production observed among program beneficiaries were not large enough to increase sales revenues and were thus mainly used for subsistence. It may also imply that market access was a challenge, a perspective which is supported by the small number of treatment villages that had a market for selling agricultural products.

Assessment of the effects of OPELIP on various well-being indicators (monthly income per adult equivalent, consumption expenditures per adult equivalent, and asset accumulation) and livelihood diversification showed that the program increased monthly income per adult equivalent by 50.9% and reduced the probability of 'income poverty' by 6.5 percentage points. Being in the program also increased monthly consumption expenditures per adult equivalent by 11.9% and reduced the probability of being 'consumption poor' by 8.8 percentage points. These results show that OPELIP reduces poverty regardless of whether poverty is measured in terms of income or consumption expenditures, which underlines the robustness of the results. Income growth can also increase consumption expenditures and reduce poverty measured in terms of consumption expenditures. OPELIP had no effect on the asset index, but it reduced the probability of being 'asset poor' by 7.4 percentage points, after controlling for confounding factors. These are important findings given that the overall goal of OPELIP is to reduce poverty and improve the wellbeing of program beneficiaries. In this case, poverty reduction mainly occurs through income growth, which increases consumption expenditures and asset accumulation. There is no evidence that the program increased (the number of) livelihood income opportunities or economic resilience as the effects of OPELIP on income diversity were insignificant.

We also examined the effects of OPELIP on a wide range of food and nutrition security indicators, including food expenditures per adult equivalent, FIES, HDDS, child and adult

anthropometric indicators, and dietary quality indicators (MDD for children, MDD for men, and MDD for women). While there were positive and significant effects of OPELIP on food expenditures per adult equivalent (a proxy for food security), the effects of OPELIP on other indicators were statistically insignificant. Although OPELIP contributed to an increase in the total value of production and incomes per adult equivalent, the growth in production and income did not translate into improved dietary diversity, dietary quality, reduced FIES, or improved adult or child anthropometry. This is contrary to the program's objectives in Component 2.

Lastly, we evaluated the effect of OPELIP on other indicators, namely, land tenure (having a land title certificate), agricultural practices (slash and burn contour farming and regular/settled agriculture), child school enrolment, and women's empowerment. The results showed that being in the program significantly increased the probability of having a plot/land title certificate by 5.9 percentage points and contributed to improved agricultural practices (reduced the probability of practicing slash and burn by 4.0 percentage points, and increased the probability of practicing regular or settled agriculture by 4.7 percentage points). However, there was no evidence to suggest that OPELIP improved women empowerment and child school enrollment as the results for these indicators were statistically insignificant.

While we expected OPELIP to have positive effects on beneficiary households (such as increased adoption of yield-enhancing inputs and nutrient-dense crop varieties that would help improve dietary diversity and dietary quality; improved agricultural productivity; increased income; improved women's empowerment and gender balance; improved child nutrition; and improved food and nutrition security) the results do not show significant effects for most of the indicators. Positive effects are only observed for chemical fertilizer use, value of production, monthly per adult equivalent incomes, consumption expenditures, land title ownership, selected agricultural practices, and poverty reduction. Part of the reasons for the insignificant effects of OPELIP could be delays in program implementation or slow uptake of certain OPELIP interventions by program beneficiaries, such as adoption of yield or soil fertility improvement innovations, adoption of kitchen gardens to promote diverse and quality diets, and women's empowerment training.

Against this background, we conclude that OPELIP is contributing to improved value of production, monthly per adult equivalent incomes, and consumption expenditures, and is also reducing poverty. However, in the next phase of OPELIP, more attention should still be given to the implementation of the program's first three components (namely, (1) community empowerment; (2) natural resource management, food and nutrition security, and livelihood improvement; and (3) community infrastructure and drudgery reduction) since the effects of the program on most of the indicators under the program's components were insignificant. We provide more detailed recommendations within these components below.

Since there is no evidence to suggest that the program has had an effect on the adoption and use of other soil fertility and yield-enhancing inputs apart from increasing the likelihood of fertilizer adoption among program beneficiaries, and descriptive results showed that sample farms (including those of treated households) are less diverse (planting only one crop per season, on average), we recommend that under Component 2, the PMU should focus on: (i) promoting the adoption of more productive varieties of various nutrient-dense and climate-resilient crops that are traditionally cultivated by the PTVGs and promote the production of seeds of promising varieties to supply village seed banks; (ii) providing training on good

agricultural practices, improved soil fertility, and soil health management; and (iii) developing grain banks and local market access infrastructure (IFAD, 2014). Implementation of some of these activities has either been delayed or has not started at all. For instance, only 11% of the targeted households had been provided with facilities for crop development and 17% of the planned kitchen gardens had been supported by the program as of August 2020 (IFAD, 2020b). Yet, these activities have the potential to enhance production and provide a marketable surplus to help beneficiary households increase their sales revenues, which have not increased so far under the program. Moreover, they could also improve household food and nutrition security.

For increased sales to be achieved, the marketable surplus resulting from increased production would need to be transported to the market. Village-level data did not provide any suggestive evidence that markets for selling agricultural products had increased or that the distance to the nearest markets had reduced in program villages. In fact, village-level data showed that only 3.5% of treatment villages had a market for selling agricultural products. Administrative data shows that out of the targeted 876 market facilities to be constructed or rehabilitated under the program, only two had been constructed as of August 2020, which means that the program is delayed on this activity (IFAD, 2020b). Hence, we recommend the construction or rehabilitation of planned markets, construction of small market yards and aggregation centers, provision of facilities for food and non-timber product processing units, including small rice hullers, and upgrading village link roads as envisioned under Component 3. This would promote market access, reduce drudgery, and boost farm and off-farm incomes.

As mentioned, the results showed that OPELIP has not had significant effects on food security, dietary quality, and anthropometric indicators for adults or children. This is partly due to delayed implementation or slow uptake of interventions aimed at improving food and nutrition security. For instance, there were delays in implementation of kitchen gardens. None of the targeted households had increased cereals, millets, oilseeds, or pulses production (IFAD, 2020b). In the next phase, the implementation should focus on promoting and training beneficiaries on kitchen gardens, crop diversification, nutrient-rich local crops (pulses, millets, and vegetables), nutritious food items (meat, eggs, fish, milk, and milk products), fruits and spices (mango, cashew, litchi, lime, orange, turmeric, ginger, chilies, etc.), and livelihood improvement activities (poultry and goat rearing).

The results also showed no evidence of improved women's empowerment by the program, perhaps due to low women's participation in women empowerment activities. As of August 2020, the participation rate in women's empowerment activities (e.g., gender training and financial literacy through SHGs) was just 28% (IFAD, 2020b). Hence, we recommend that under Component 1, the PMU should focus on how to engage women to increase their participation in SHGs and women's empowerment activities (such as gender training and financial literacy) to improve their empowerment as well as the gender balance. Also under Component 1, the PMU should aim to improve the performance ratings of village development associations (VDAs), which seem to be declining based on the village-level results. Strengthened VDAs will improve the planning and execution of interventions based on the needs of the community.

References

Alkire, S., Meinzen-Dick, R., Peterman, A., Agnes Quisumbing, Seymour, G., & Vaz, A. (2013). The Women's Empowerment in Agriculture Index. *World Development*, *52*(0), 71–91.

Andersen, L. E., & Cardona, M. (2013). *Building resilience against adverse shocks: What are the determinants of vulnerability and resilience?* (No. 02/2013). Development Research Working Paper Series.

Angelucci, M., & Di Maro, V. (2016). Programme evaluation and spillover effects. *Journal of Development Effectiveness*, *8*(1), 22–43.

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics. Princeton university press.

Ballard, T. J., Kepple, A. W., & Cafiero, C. (2013). *The food insecurity experience scale: development of a global standard for monitoring hunger worldwide*. Rome: FAO, 61.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., … Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, *348*(6236), 1260799.

Bird, F. A., Pradhan, A., Bhavani, R. V., & Dangour, A. D. (2019). Interventions in agriculture for nutrition outcomes: A systematic review focused on South Asia. *Food Policy*, *82*, 39–49.

Cahyadi, E. R., & Waibel, H. (2016). Contract farming and vulnerability to poverty among oil palm smallholders in Indonesia. *The Journal of Development Studies*, *52*(5), 681–695.

Cameron, A. Colin, and Pravin K. Trivedi. *Microeconometrics: methods and applications*. Cambridge university press, 2005.

Choudhury, S., Shankar, B., Aleksandrowicz, L., Tak, M., Green, R., Harris, F., ... & Dangour, A. (2020). What underlies inadequate and unequal fruit and vegetable consumption in India? An exploratory analysis. *Global Food Security*, *24*, 100332.

CNNS. (2019). *Ministry of Health and Family Welfare, Government of India, UNICEF, and Population Council. 2019*. "Comprehensive National Nutrition Survey: 2016–2018." New Delhi. https://www.unicef.org/india/media/2646/file/CNNS-report.pdf

Davis, K., Nkonya, E., Kato, E., Mekonnen, D. A., Odendo, M., Miiro, R., & Nkuba, J. (2012). Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa. *World Development*, 40(2), 402–413.

de Brauw, A., Gilligan, D. O., Hoddinott, J., & Roy, S. (2014). The Impact of Bolsa Família on Women's Decision-Making Power. *World Development*, *59*, 487–504.

Dillon, A. (2011). The effect of irrigation on poverty reduction, asset accumulation, and informal insurance: Evidence from Northern Mali. *World Development*, *39*(12), 2165–2175.

FAO & FHI 360. (2016). Minimum dietary diversity for women: a guide for measurement 2016. Rome. Available from: http://www.fao.org/3/a-i5486e.pdf (accessed February 2022).

FAO (2015). *Livelihood diversification and vulnerability to poverty in rural Malawi*, by Solomon Asfaw, Nancy McCarthy, Adriana Paolantonio, Romina Cavatassi, Mulubrhan Amare and Leslie Lipper. ESA Working Paper No. 15-02. Rome.

FAO. (2021). *Minimum dietary diversity for women*. Food and Agriculture Organization of the United Nations. Rome. https://doi.org/10.4060/cb3434en

Fiorella, K. J., Chen, R. L., Milner, E. M., & Fernald, L. C. H. (2016). Agricultural interventions for improved nutrition: A review of livelihood and environmental dimensions. *Global Food Security*, *8*, 39–47.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice, Second Edition*. World Bank Publications.

Greene, W. H. (2012). Econometric Analysis, 7th edition. Upper Saddle River, NJ: Prentice Hall.

Gupta, S., Brazier, A. K. M., & Lowe, N. M. (2020). Zinc deficiency in low-and middle-income countries: prevalence and approaches for mitigation. *Journal of Human Nutrition and Dietetics*, 33(5), 624–643.

Hertel, T., Elouafi, I., Tanticharoen, M., & Ewert, F. (2021). Diversification for enhanced food systems resilience. *Nature Food*, 2(11), 832-834.

IFAD (2014, August). *Final Project Design Report: Odisha PTG Empowerment and Livelihoods Improvement Programme*. Asia and the Pacific Division, Program Management Department, International Fund for Agricultural Development (IFAD).

IFAD (2017). *Impact Assessment Baseline Plan. Odisha Particularly Vulnerable Tribal Group Empowerment and Livelihood Improvement Programme (OPELIP)*. Rome: IFAD.

IFAD (2020a, December). *Implementation Support Report: Odisha PTG Empowerment and Livelihoods Improvement Programme.* Asia and the Pacific Division, Program Management Department, International Fund for Agricultural Development (IFAD).

IFAD (2020b, December). *Supervision Report: Odisha Particularly Vulnerable Tribal Groups Empowerment and Livelihoods.* Asia and the Pacific Division, Program Management Department, International Fund for Agricultural Development (IFAD).

Jodlowski, M., Winter-Nelson, A., Baylis, K., & Goldsmith, P. D. (2016). Milk in the Data: Food Security Impacts from a Livestock Field Experiment in Zambia. *World Development*, *77*, 99–114.

Kafle, K., Jolliffe, D., & Winter-Nelson, A. (2018). Do different types of assets have differential effects on child education? Evidence from Tanzania. *World Development*, *109*, 14–28.

Kafle, K., Winter-Nelson, A., & Michelson, H. (2016). Does she have a say? The impact of livestock transfer and associated training on women's empowerment: Evidence from Zambia. In *2016 Annual Meeting, July 31-August 2, 2016, Boston, Massachusetts*. Agricultural and Applied Economics Association.

Kennedy, G., Ballard, T., & Dop, M. C. (2010). *Guidelines for measuring household and individual dietary diversity*. Food and Agriculture Organization of the United Nations.

Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations [the interpretation of dummy variables in semilogarithmic equations]. *American Economic Review*, *71*(4), 801–801.

Lawry, S., Samii, C., Hall, R., Leopold, A., Hornby, D., & Mtero, F. (2017). The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review. *Journal of Development Effectiveness*, *9*(1), 61–81.

Leroy, J. L., Ruel, M., Frongillo, E. A., Harris, J., & Ballard, T. J. (2015). Measuring the food access dimension of food security: a critical review and mapping of indicators. *Food and nutrition bulletin*, *36*(2), 167–195.

Liverpool-Tasie, L. S. O., & Winter-Nelson, A. (2011). Asset versus consumption poverty and poverty dynamics in rural Ethiopia. *Agricultural Economics*, *42*(2), 221–233.

Martey, E., Etwire, P. M., & Mockshell, J. (2021). Climate-smart cowpea adoption and welfare effects of comprehensive agricultural training programs. *Technology in Society*, 64, 101468.

Muriithi, B. W., & Matz, J. A. (2015). Welfare effects of vegetable commercialization: Evidence from smallholder producers in Kenya. *Food policy*, *50*, 80–91.

Nakano, Y., Tsusaka, T. W., Aida, T., & Pede, V. O. (2018). Is farmer-to-farmer extension effective? The impact of training on technology adoption and rice farming productivity in Tanzania. *World Development*, *105*, 336–351.

NFHS 5. (2019-21). *National Family Health Survey 5, 2019-21*. Government of India, Ministry of Health and Family Welfare. http://rchiips.org/nfhs/NFHS-5_FCTS/Odisha.pdf

Ogutu, S. O., Fongar, A., Gödecke, T., Jäckering, L., Mwololo, H., Njuguna, M., ... & Qaim, M. (2020b). How to make farming and agricultural extension more nutrition-sensitive: evidence from a randomised controlled trial in Kenya. *European Review of Agricultural Economics*, *47*(1), 95–118.

Ogutu, S. O., Ochieng, D. O., & Qaim, M. (2020a). Supermarket contracts and smallholder farmers: Implications for income and multidimensional poverty. *Food Policy*, *95*, 101940.

OTELP, 2016 OTELP. (2016). *Main Completion Report: Odisha Tribal Empowerment and Livelihoods Programme. OTELP Programme Support Unit, ST & SC Development Department*. Government of Odisha. http://otelp.org/Downloads/Project-Completion-Report-2016.pdf

Pandey, V. L., Mahendra Dev, S., & Jayachandran, U. (2016). Impact of agricultural interventions on the nutritional status in South Asia: A review. *Food Policy*, *62*, 28–40. https://doi.org/10.1016/j.foodpol.2016.05.002

Rao, Y. G., Pattanik, S. C., Barada, H. B., & Patra, S. (2020). *Base Line Survey of PVTGs in Odisha.*

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Ruel, M. T., Quisumbing, A. R., & Balagamwala, M. (2018). Nutrition-sensitive agriculture: what have we learned so far?. *Global Food Security*, *17*, 128–153.

Sahoo, P. K., Biswal, S., Kumar, H., & Powell, M. (2022). Urban to rural COVID-19 progression in India: The role of massive migration and the challenge to India's traditional labour force policies. *The International Journal of Health Planning and Management*, *37*(1), 528–535.

Sharma, I. K., Di Prima, S., Essink, D., & Broerse, J. E. (2021). Nutrition-sensitive agriculture: a systematic review of impact pathways to nutrition outcomes. *Advances in Nutrition*, *12*(1), 251–275.

United Nations (2004). World youth report 2003. Retrieved from http://www. un.org/esa/socdev/unyin/documents/ worldyouthreport.pdf (accessed 19.07.2022)

Wan, J., Li, R., Wang, W., Liu, Z., & Chen, B. (2016). Income diversification: A strategy for rural region risk management. *Sustainability*, 8(10), 1064.

WHO. (2011). Waist circumference and waist-hip ratio: report of a WHO expert consultation, Geneva, 8-11 December 2008. World Health Organization, Geneva.

WHO Multicentre Growth Reference Study Group (Ed.). (2006). WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. World Health Organization, Geneva.

WHO & UNICEF. (2021). *Indicators for assessing infant and young child feeding practices: definitions and measurement methods*. Geneva: World Health Organization and the United Nations Children's Fund (UNICEF), 2021. Licence: CC BYNC-SA 3.0 IGO; https://creativecommons.org/licenses/by-nc-sa/3.0/igo.

WHO. (2022). *Body mass index*. World Health Organization. https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi (accessed February 2022)

Winters, P., Salazar, L., & Maffioli, A. (2010). *Designing Impact Evaluations for Agricultural Projects* (Technical Notes No. IDB-TN-198). Inter-American Development Bank. World Bank. (2007). *Data for Impact Evaluation* (Doing Impact Evaluation No. N. 6). Washington, DC: The World Bank.

World Bank, T. (2011). *Impact evaluations in agriculture: an assessment of the evidence* (No. 72693) (pp. 1–104). Washington, DC: Independent Evaluation Group (IEG), World Bank.

World Bank. (2007). *Data for impact evaluation (English)*. Doing impact evaluation series; no. 6 Washington, D.C: World Bank Group. http://documents.worldbank.org/curated/en/332891468313760995/Data-for-impact evaluation.

World Bank. (2022). Consumer price index for India. Available at https://data.worldbank.org/indicator/FP.CPI.TOTL?locations=IN (accessed February 2022).

Appendix

Table A1 IFAD 11 indicators

Tier II Development Impact Indicators	IFAD Target	SDG	RMF Code
Household income	EG	1.2 and 2.3	2.1.1
Asset ownership	EG	1.2 and 2.3	2.1.1
Consumption			
Value of production	SO1	2.3.2	2.1.2
Value of market sales	SO2	2.3	2.1.3
Resilience	SO3	1.5	2.1.4
Nutrition/Food security	MT	2.1	2.1.5
Gender	MT		
Youth	MT		
Persons with a disability	MT		
Nutrition	MT		
Climate change	MT		
Land Tenure	Additional Core		
Job Creation	Additional Core		

Table A2 Mean Comparison of child anthropometric (0 to 59 months of age) indicators by treatment, control and tribal group in 2021

2021					
	Total	Treatment	Control	PVTG	Non-PVTG
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Mild prevalence					
Stunted (<-1 SD) (%)	0.722	0.701	0.745	0.740	0.711
	(0.448)	(0.458)	(0.437)	(0.439)	(0.454)
Underweight (<-1 SD) (%)	0.717	0.705	0.730	0.760	0.690*
	(0.451)	(0.457)	(0.445)	(0.428)	(0.463)
Wasted (<-1 SD) (%)	0.492	0.455*	0.533	0.534	0.466
	(0.500)	(0.499)	(0.500)	(0.500)	(0.500)
Overweight (>+1 SD) (%)	0.0768	0.0660	0.0888	0.0913	0.0678
	(0.266)	(0.249)	(0.285)	(0.289)	(0.252)
Standard prevalence					
Stunted (<-2 SD) (%)	0.446	0.427	0.467	0.481	0.425
	(0.498)	(0.496)	(0.500)	(0.501)	(0.495)
Underweight (<-2 SD) (%)	0.400	0.361**	0.444	0.433	0.381
	(0.490)	(0.481)	(0.498)	(0.497)	(0.486)
Wasted (<-2 SD) (%)	0.148	0.139	0.158	0.178	0.130
	(0.356)	(0.346)	(0.366)	(0.383)	(0.337)
Overweight (>+2 SD) (%)	0.0238	0.0104**	0.0386	0.0240	0.0236
	(0.152)	(0.102)	(0.193)	(0.154)	(0.152)
Severe prevalence					
Stunted (<-3 SD) (%)	0.207	0.205	0.208	0.240	0.186
	(0.405)	(0.404)	(0.407)	(0.428)	(0.390)
Underweight (<-3 SD) (%)	0.113	0.125	0.100	0.144	0.0944*
	(0.317)	(0.331)	(0.301)	(0.352)	(0.293)
Wasted (<-3 SD) (%)	0.0347	0.0347	0.0347	0.0433	0.0295
	(0.183)	(0.183)	(0.183)	(0.204)	(0.169)
Observations	547	288	259	208	339

Note: Point estimates are sample means. Standard deviations are displayed in parentheses at the columns total, treatment and control, while p-values are displayed in parentheses at the column mean differences.
*****A t-test is used to compare the mean differences between PVTG and non-PVTG in 2021. Level of significance $^{*}p< 0.10$, $^{**}p< 0.05$, $^{***}p< 0.01$.