

Improving Incomes of Coffee Farmers (IICoF) in Jimma, Ethiopia

Baseline report

Gashaw T. Abate, Kalle Hirvonen, Belay Mohammed

November 2024

Submitted to: HereWeGrow



A world free of hunger and malnutrition

IFPRI is a CGIAR Research Center

1201 Eye Street, NW, Washington, DC 20005 USA |

T. +1-202-862-5600 | F. +1-202-862-5606 | Email:

ifpri@cgiar.org | www.ifpri.org | www.ifpri.info

Table of Contents

Executive summary	6
1. Introduction	9
2. The interventions	10
2.1. Theory of change	11
3. Methodology.....	11
3.1. Evaluation questions.....	11
3.2. Evaluation design	12
3.3. Primary and secondary outcomes	14
3.4. Sample and statistical power	16
3.5. Data and timeline.....	17
4. Results and discussions.....	20
4.1. Household demographics and socio-economic characteristics	20
4.2. Coffee and maize production, sales, knowledge and adoption of recommended practices.....	26
4.3. Production of other crops.....	36
4.4. Livestock ownership.....	36
4.5. Incomes.....	37
4.6. Access to loans.....	40
4.7. Consumption.....	43
4.8. Food security.....	46
4.9. Women’s self-help groups	48
5. Summary and conclusions	52
References	54
Appendix	55
Appendix A: Randomization and sampling protocol	55
Appendix B: Baseline balance	61
Appendix C: Woreda level results.....	65
Appendix D: Supplementary figures on household welfare outcomes.....	68
Appendix E: Supplementary figures on adoption of coffee best practices	69
Appendix F: Supplementary figure on locations of coffee sold.....	70
Appendix G: Supplementary table on maize and coffee knowledge and adoption rules.....	71

List of Tables

Table 1. Primary outcome indicators.....	15
Table 2. Secondary outcome indicators	16
Table 3. Timing of randomization, surveys, and intervention	18
Table 4. Farm area, by study arm	25
Table 5. Household level coffee production, sales, and coffee revenue during the 2023 (2015/16 E.C.) production season	27
Table 6. Knowledge of recommended coffee farming practices.....	28
Table 7. Household level maize production, sales, and maize revenue during the 2023 (2015/16 E.C.) production season	34
Table 8. Knowledge of recommended maize farming practices	34
Table 9. Monetary loan, inputs credit and goods credit in the last 12 months	41
Table 10. Household consumption and poverty indicators.....	44
Table 11. Food Insecurity Experience Scale questions	47
Table 12. Household Diet Diversity Score (HDDS)	48
Table 13. Women’s self-help group activities.....	49

List of Figures

Figure 1. Evaluation design	14
Figure 2. Housing characteristics	21
Figure 3. Average walking distances to roads, markets, agricultural offices and financial institutions	22
Figure 4. Ownership of durable assets	23
Figure 5. Distribution of the asset index values	24
Figure 6. Farm size categories, total and coffee and maize plots	25
Figure 7. Recommended coffee farming practices: distribution of knowledge score	28
Figure 8. Stumping intensity	29
Figure 9. Organic fertilizer uses	30
Figure 10. Weeding methods.....	31
Figure 11. Shade level	32
Figure 13. Main disease problems	32
Figure 14. Summary of best practice adoption	33
Figure 15. Recommended maize farming practices: distribution of the knowledge score.....	35
Figure 16. Adoption of recommended maize farming practices	35
Figure 17. Other crops cultivated by the farmers.....	36
Figure 18. Percent of households owning different livestock species	37
Figure 19. Median number of livestock owned (conditional on owning), by livestock type	37
Figure 20. Mean and median total per capita incomes, by income quintile.....	38
Figure 21. Mean total incomes, by income source and income quintile	39
Figure 22. Mean crop incomes, by crop and income quintile	39
Figure 23: Source of loan, purpose, reasons for delayed payment, and potential lenders	42
Figure 24. Distribution of household daily per capita consumption in 2017 PPP	44
Figure 25. Risk of extreme poverty, by household size	45
Figure 26. Risk of extreme poverty, by land size	46
Figure 27. Distribution of the Food Insecurity Experience Scale (FIES) scores.....	47
Figure 28. Percent of women belonging to a women’s self-help group, by woreda	49
Figure 29. Decision making within households across different domains	50
Figure 30. Degree to which women feel they can make their own personal decision	51

Acronyms

AIC	Access to Inputs and Input credit
ANCOVA	Analysis of covariance
CAPI	Computer Assisted Personal Interview
CPI	Consumer Price Index
DA	Development Agent
DE	Digital Extension
FIES	Food Insecurity Experience Scale
FTC	Farmer Training Center
HDDS	Household Diet Diversity Score
HWG	HereWeGrow
ICC	Intra-cluster Correlation coefficient
ICP	International Comparison Program
IFPRI	International Food Policy Research Institute
IICoF	Improving Incomes of Coffee Farmers
IPDM	Integrated Pest and Disease Management
IRB	Institutional Review Board
MDE	Minimum Detectable Effect
PSNP	Productive Safety Net Program
RCT	Randomized Controlled Trial
TLU	Tropical Livestock Units
WSHG	Women's Self-Help Groups

Acknowledgments

This baseline study received financial support from HereWeGrow. We are thankful to our colleagues at Digital Green for their thoughtful collaboration during the household listing exercise and baseline data collection. We are grateful to the invaluable inputs of HWG colleagues, which improved the report considerably. We acknowledge Greenpact and their survey staff for collecting the baseline data. We are especially grateful for the time and willingness of the many households who participated in the baseline survey. All errors are the sole responsibility of the authors. The opinions expressed here belong to the authors, and do not necessarily reflect those of IFPRI, Digital Green, or HWG.

Executive summary

Limited adoption of agricultural technologies and best practices remain the biggest barrier to raising crop yields in Ethiopia. Several social, economic, and behavioral factors contribute to low adoption rates, including difficulties obtaining inputs, liquidity and credit constraints, risks and uncertainty, lack of clear information, and a tendency to focus on short-term gains. To address these challenges and increase smallholder productivity, HereWeGrow, in collaboration with Digital Green, is implementing a project that includes a bundle of interventions. The interventions targets smallholder farmers residing in coffee-producing areas within the Jimma zone of Ethiopia. In this context, farmers primarily engage in the cultivation of cash crops, with coffee being the most important crop both in terms of area coverage and income. Additionally, alongside coffee, farmers cultivate maize primarily for home consumption.

As part of the project, IFPRI is undertaking a rigorous impact evaluation to generate valuable insights and lessons that can be applied to scaling up similar efforts in Ethiopia and other low-income settings. Specifically, the evaluation relies on a randomized controlled trial (RCT) to assess the causal effect of the interventions on the adoption of good agricultural practices, farm productivity, and household income. The baseline survey for this impact evaluation was conducted in December 2023 and January 2024 and this report provides baseline results based on 2,995 sample households from 750 development groups (clusters) in 75 target/study kebeles.

Household characteristics: The average sample household include more than 5 members and is headed by a 45-year-old household head, who is typically male (in 91 percent of the households) and illiterate (in 43 percent of the households). Most households reportedly have limited access to amenities (e.g., electricity, piped water), even though they reside within less than an hour walking distance to all-weather road, weekly market and primary coffee market, coffee washing stations, agricultural cooperatives, and farmer training center/DA office.

Agricultural land and assets: The average household in our sample operates 0.97 ha of agricultural land, of which about half (0.51 ha) is allocated to coffee farming and 0.27 ha to maize farming and 0.19 ha to other crops. While most households own durable assets such as traditional farm equipment, solar panel, mobile phone, and various types of furniture, ownership of more valuable assets such as motorbike, TV, water pump, tractor, harvester, etc. is limited. Households in our sample own 1.2 tropical livestock units (TLU), on average.

Coffee and maize production and sales: Sample households produced 157 kg of green beans during the 2023 season (246 kg of cherry and 216 kg of jenfel). The green bean yield was 348 kg/ha. By the time of the interview, the mean amount of cherry coffee sold by the farmers was 160 kg, and jenfel coffee was 66 kg. These sales translate to mean revenue of 4,690 birr from selling cherries and 4,620 birr from selling jenfel, totaling 9,310 birr. More than 70 percent of the farmers in our sample produced maize and the average maize farmer reported producing 694 kg of maize, and the average yield was about 2,000 kg per hectare. Most of the maize is produced for home consumption, as only 13 percent of the farmers reported selling or planning to sell the maize they had harvested.

Knowledge and adoption of best/recommended practices: While sample households showed a high degree of understanding of recommended practices for shade management and stumping methods, they have relatively limited knowledge on other coffee best practices. Out of a maximum of seven points on the knowledge quiz, the average farmer in our sample scored 4.2 points. Meanwhile, more than 82 percent of the farmers practiced integrated disease management (IDM). Other recommended best practices, such as planting new trees and soil fertility were followed/practiced by about 50 percent and 20 percent of the farmers, respectively. Only about 10 percent of the farmers reported to have stumped (rejuvenated) their coffee trees in the last three years (~4 percent each year). Similarly, while farmers showed a good grasp of some of the recommended maize production practices such as ploughing frequency, NPS/DAP fertilizer application and Fall Armyworm control methods, they lack the basic understanding on UREA application, inter-cultivation and row planting techniques. Moreover, most farmers did not apply the recommended per-hectare amount of NPS or urea fertilizers on maize farms.

Income: The mean total per capital annual income in the sample is 14,771 birr (median 9,968 birrs). Looking at income by quintiles, the median annual per capita household income in the bottom quintile (i.e., the poorest 20%) is 3,510 birrs, while at the top quintile (the richest 20 %) it is 27,693 birrs. In each quintile, crop incomes form the lion's share of the total income and coffee is the most important source of total crop income.

Access to loans: Only one-third of the sample households applied for loan over the 12 months prior to the survey and the remainder (two-third) didn't apply because of reasons such as not needing a loan (73 percent of those who didn't apply), expecting to be rejected (15 percent), lacking collateral (5 percent), and thinking that the interest rate was too high (3 percent). While most of the households that applied did obtain the loan, a third of them were unable to make one or more payments, typically because the payments were due at a time when households had less income than usual. More than 90 percent of these loans were obtained from friends and relatives / informal sources. Obtaining input on credit is rare in this context with only one percent of the households reporting having obtained agricultural inputs on credit.

Consumption: The mean daily household per capita food consumption-expenditures in this sample is 72.42 birr and the mean daily per capita non-food expenditures is 11.26 birr. In aggregate, the average household in our sample spends 86 percent of its total budget on food, indicative of the high level of deprivation in this sample of households. The mean total per capita consumption expenditure (83.68 birr per day) is equivalent to 2.33 \$PPP, which is only marginally higher than the international extreme poverty line: \$2.15 \$PPP. In fact, more than half (60 percent) of the individuals originate from households for which the daily per capita consumption is below 2.15 \$PPP. Many households in our sample are located just below or just above the poverty line, indicating that even relatively small changes in consumption levels may induce large shifts in poverty headcount estimates.

Food security: Results from the Food Insecurity Experience Scale (FIES) indicate that 24 percent of the households were food secure, 35 percent mildly food insecure, 28 percent moderately food insecure and 12.5 percent severely food insecure. In terms of food/dietary diversity, the average household in our sample consumed 7.6 food groups out of 12. Almost all households

consumed Cereals, Vegetables, Pulses, Oils and Miscellaneous foods, while the consumption of animal source foods was relatively less common.

Women Self-Help Groups (WSHGs): About 9 percent of the women in the sample reported belonging to a women's self-help group (WSHG) and most of these groups operated in the Shabe Sombo woreda, where one-third of the sampled women belonged to a WSHG. A typical WSHG has about 30 women, and about two-thirds of the groups organize training for their members on agriculture (44 percent), business management (56 percent), financial management (71 percent). The self-help groups also provide services such as savings (70 percent), loan/credit (39 percent), and income generating opportunities (3 percent) to their members.

1. Introduction

Despite Ethiopia's significant efforts to boost agricultural output through the promotion of modern inputs and farming methods over the past 30 years, smallholder farmers have not fully embraced these modern practices (Berhane, Abate, and Wolle 2022). This limited adoption remains the biggest barrier to raising crop yields in the country (Berhane, Abate, & Wolle, 2022). Several social, economic, and behavioral factors contribute to low adoption rates, including difficulties obtaining inputs, liquidity and credit constraints, risks and uncertainty, lack of clear information, and a tendency to focus on short-term gains.

A case in point is coffee, a crop that employs more than 15 million people along the value chain in Ethiopia and generates one-third of the country's export earnings. Despite being Africa's largest coffee producer and exporter, Ethiopia faces significantly lower productivity compared to other major coffee-producing countries. For example, Arabica coffee yields in Brazil are more than four times higher than in Ethiopia (1,600 kg green/ha in Brazil versus 378 kg green/ha in Ethiopia). This productivity gap primarily results from aging coffee trees and inadequate rejuvenation practices among Ethiopian smallholder farmers that together account for 90 percent of total coffee production in the country. Recent World Bank (2021) estimates suggest that 80 percent of Ethiopia's coffee trees remain unpruned or unstumped, leading to reduced productivity, further compounded by limited knowledge of good agronomic practices, and poor access to inputs and credit. Low and declining productivity directly threatens millions of smallholder coffee farmers and other value chain participants reliant on coffee for their livelihoods. Indirectly, low coffee productivity contributes to Ethiopia's chronic foreign exchange scarcity that risks future growth prospects.

To address these challenges and increase smallholder productivity, HereWeGrow, partnering with Digital Green, is launching a project in Jimma Zone in Ethiopia's Oromia region. The intervention targets smallholder farmers residing in coffee-producing areas within the Jimma zone of Ethiopia. In this context, farmers primarily engage in the cultivation of cash crops, with coffee being the most important crop both in terms of area and income. Additionally, alongside coffee, farmers cultivate maize for subsistence purposes. Coffee farmers in Jimma are largely poor and food insecure. In a 2022-survey in just one Jimma woreda (district), more than 90 percent of the 950 coffee farming households reported some level of food insecurity with 60 percent reporting moderate or severe food insecurity (Abate, Hirvonen, and Mohammed 2022).

The project aims to improve the income and livelihoods of smallholder farmers who rely on coffee as their primary source of income, while also recognizing the importance of maize in ensuring food security. The bundled intervention attempts to address multiple technology adoption constraints in coffee and maize farming with the goal of improving farm productivity and household incomes. This evaluation aims to generate valuable insights and lessons that can be applied to scaling up similar efforts in Ethiopia and other low-income settings. Specifically, the evaluation employs a rigorous impact assessment based on an experimental design with multiple treatment arms assessing the overall impact of the project and the marginal impact of key

intervention components. The evaluation will mainly rely on a randomized controlled trial (RCT) to assess the causal effect of the interventions on the adoption of good agricultural practices, productivity, and income. The evaluation will also assess potential synergies between the interventions, provide insights into the pathways that help explain the realized impacts, and potential heterogeneity of impacts along economic and social dimensions.

The baseline survey for this impact evaluation was conducted in December 2023 and January 2024. This baseline report provides an overview of the evaluation design, the survey conducted, and the key findings. These findings allow us to characterize the sample and verify the balance on observable characteristics across experimental arms.

2. The interventions

A focus group discussion conducted in 2022 in Jimma by Abate, Hirvonen, and Mohammed (2021) highlighted coffee as the predominant crop among farmers in Jimma, significantly contributing to both the cultivated land area and household income. While many farmers also cultivated maize, coffee was estimated to constitute approximately 50 percent of the total household income. Moreover, the majority (75%) of the coffee trees were over 20 years old. Participants in the focus discussion groups highlighted several barriers to rejuvenation, including lack of knowledge and appropriate tools for stumping coffee trees, a shortage of coffee seedlings, and the effort required to maintain them properly – especially on farms located far from the dwelling. Additionally, some participants emphasized that rejuvenating coffee farms meant a loss of income source for a couple of years, which many farmers could not afford due to coffee farming constituting a significant portion of their income. Better access to coffee production inputs, market linkages, and coffee agronomy training were ranked among the top three priorities by the farmers. These factors were seen as critical for overcoming barriers to coffee farm renewal and improving coffee cultivation practices.

The bundled intervention package in this project has been carefully designed to address these multiple constraints faced by the farmers in this context. To increase the adoption of coffee tree rejuvenation and other good agricultural practices, Digital Green will train farmers using video and interactive voice response technologies and build capacity of government extension workers to provide extensions using these technologies. To improve financial inclusion and facilitate productive investments (e.g., inputs), the project will partner with a local bank to improve access to affordable credit for farm households. Moreover, to ensure the availability of inputs, the project will facilitate timely availability of inputs and tools, provide financial and technical support to government-managed nurseries, and deliver training and improved seeds to coffee nurseries operated by households. To mitigate potential negative externalities stemming from the crowding out of inputs, as discussed by Duflo et al. (2023), the input component will be distributed across the entire project area, encompassing both the intervention arms and the control arm situated within the project area. The project's gender component establishes women's self-help groups (WSHG) and uses these groups as a platform to deliver (video-based)

training on group governance, conflict management, financial literacy, business development, and entrepreneurial skills. In addition, the groups will be provided with seed capital to strengthen group lending capacity to develop new and strengthen existing livelihood activities.¹

2.1. Theory of change

The main theory of change for this project is that the digital extension improves farmer's knowledge about good agronomic practices. More knowledgeable coffee and maize farmers are more likely to adopt good agronomic practices (e.g., stump old coffee trees; use improved seedlings or seeds; apply fertilizers and pesticides) that lead to improved yields. The likelihood of adoption is enhanced by improving farmers' access to credit and inputs.

The WSHG project component uses (video-based) training and saving and credit supports to improve women's entrepreneurial skills and financial inclusion. Increased knowledge and financial resources would permit setting up or expanding existing businesses, thus providing an opportunity to diversify and increase incomes.

The shared primary objective in all project components is to increase smallholder incomes.

3. Methodology

3.1. Evaluation questions

The primary objective of the evaluation is to quantify the *causal impact* of the three key components of the project (i.e., digital extension, input credit, and WSHG) on the adoption of good agricultural practice, yield, household income, and other relevant outcomes. Specifically, the evaluation is designed to answer the following eight evaluation questions under digital extension (DE), access to inputs and input credit (AIC), and women self-help groups (WG):

Digital extension:

- DE1: What are the take-up rates of video-based training?
- DE2: What socio-economic baseline characteristics drive participation in digital extension?
- DE3: What is the effect of digital extension on the adoption of the promoted practices in coffee and maize cultivation?

Access to inputs and input credit:

- AIC1: Are coffee and/or maize farming households in the selected project kebeles able to purchase coffee inputs (seedlings and tools) and maize inputs (fertilizer, seeds) at the time, quantity, and variety desired?

¹ The project includes a component supporting 225 honey producer households in selected woredas by providing digital extension and honey production equipment. This small pilot component is not amenable for a rigorous evaluation and therefore its impact is not assessed in this research.

- AIC2: How do credit applicants differ from non-applicants in project kebeles?
- AIC3: What are the treatment effects of input credit + digital extension on:
 - a) Adoption of best agricultural practices in coffee and maize cultivation
 - b) Input use in coffee and maize cultivation
 - c) Expenditures in coffee and maize cultivation
 - d) Coffee and maize production and productivity
 - e) Revenues (cash and in-kind gross income) from coffee and maize production
 - f) Income-generating activities of the household
 - g) Total household income
 - h) Total household consumption
 - i) Asset holdings

Women Self-Help Groups (WSHG):

- WG1: Who participates in WSHGs? How do members differ from non-members?
- WG2: How do the following outcomes develop over time for households where the woman participates in the WSHG component:
 - a) Income-generating activities of the household
 - b) Savings and loans
 - c) Total household income (disaggregated by income source)
 - d) Asset holdings

3.2. Evaluation design

To answer the abovementioned research questions and to assess the value added of combining digital extension focusing on both coffee and maize instead of only coffee, the evaluation designed a cluster randomized control trial (RCT) in the four project woredas. Cluster is defined at the development group level. The study clusters have been randomly allocated to different study arms, permitting us to identify the causal impact of the project. The evaluation relies on longitudinal data that follow sample coffee farming households for a four-year period covering pre-treatment (baseline) period during the treatment (midline), and post-treatment period (endline).

Figure 1 summarizes the evaluation design. We will first randomize the 75 project kebeles into 55 'Treatment' kebeles and 20 'Pure Control' kebeles (the green boxes in Figure 1). In the 'Treatment' kebeles we draw a random sample of 625 development groups. To understand potential synergies between the digital extension (DE), input credit (IC), and WSHG project components, these 625 development groups were randomized into 5 study arms (the yellow boxes in Figure 1), each consisting of 125 development groups:

- Study arm T1: DE on coffee
- Study arm T2: DE on coffee & maize
- Study arm T3: DE on coffee & maize + IC
- Study arm T4: DE on coffee & maize + IC + WSHG
- Study arm TC: Control

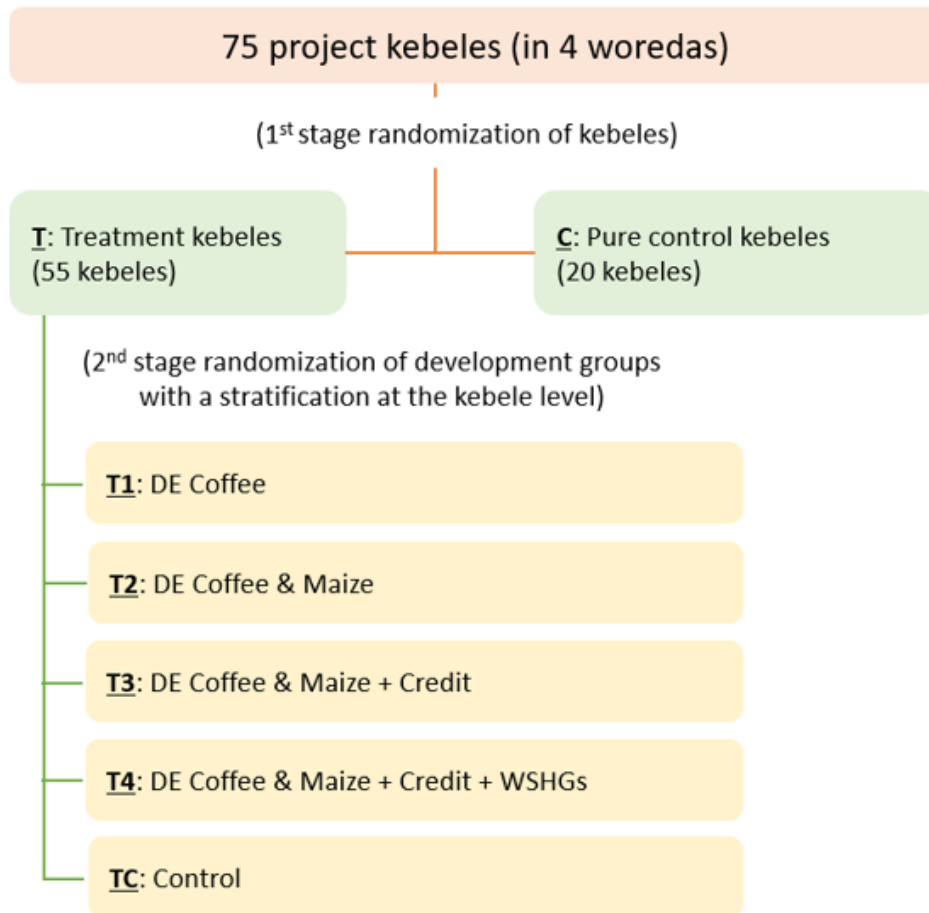
The households in arm T1 are exposed to Digital Green’s Video Extension services that focus on coffee farming.² The households in study arm T2 are exposed to video extension services that focus on both coffee and maize farming. The households in study arm T3 receive the same intervention bundle as households in study arm T2 but in addition, are offered input credit facilitated by the project. Households in study arm T4 are exposed to the digital extension on coffee and maize, receive input credit, and are offered to join women’s self-help groups.

We measure the impacts of the project components on the outcomes of interest against the Control arm (TC) that does not directly benefit from any of the main intervention components. Comparisons of the treatment arms (T1 to T4) will inform about the complementarities of the different treatment bundles. For example, a comparison of outcomes between households in T2 and households in T3 arms will allow us to quantify the value-added of bundling DE with input credit. Comparison of T1 and T2 arms will inform us of the effectiveness of combining DE messaging on coffee and maize farming against a DE that focuses only on coffee farming.

To identify and quantify possible spillover effects, we include ‘Pure Control’ kebeles from which we have drawn a random sample of 125 development groups. We perceive that there is a non-negligible risk that the farmers who are not directly exposed to the DE treatment may learn from other farmers who are exposed to the DE treatments and reside in the same kebele. Such positive spill-over effects would mean that the incomes of the households in TC also improve because of the Digital Green activities taking place in the kebeles. As a result, the comparisons between TC and other study arms would underestimate the project’s impact. If these positive spillover effects are sufficiently large, there is a real risk of an erroneous conclusion that the project did not have any impact. To measure these spillover effects, we propose to compare the outcomes between TC control households and the households in the ‘Pure Control’ kebeles. While neither of these groups will directly benefit from the project, the TC households may indirectly benefit through learning from treated households residing in the same kebeles. The proposed comparison of these two groups of control households would allow us to quantify these indirect benefits of the project.

² Development groups in T1 will receive digital extension services on maize after two production seasons.

Figure 1. Evaluation design



Note: DE = Digital Extension; DG = Development Group; WSHGs = Women self-help groups. Development groups in T1 will receive digital extension services on maize after two production seasons.

3.3. Primary and secondary outcomes

The primary outcomes (listed in Table 1) in this evaluation align with the project's theory of change, evaluating the effectiveness of digital extension in improving farmers' knowledge of good agricultural practices. If the digital extension is effective, we should first see improvements in farmers' knowledge of good agricultural practices. To measure this, the quantitative survey instruments include a quiz testing farmer's knowledge of these practices. The final variable used in the analysis is a simple count capturing the number of correct responses in this quiz.

Improved knowledge should lead farmers to adopt these good agronomic practices (i.e., stumping, coffee nutrition, weeding, etc. in the case of coffee and improved seeds, fertilizers, pesticides, etc. in the case of maize) promoted by Digital Green. We measure this with a set of simple yes/no questions on whether the farmer adopted these practices in the past 12 months. We then compute a simple count variable capturing the number of good agronomic practices

adopted by the farmer at the time of the interview. Considering the project focus on stumping to address the problem of aging/old coffee trees in Ethiopia, we analyze the impact on coffee tree stumping separately. This variable will obtain a value of 1 if the farmer has stumped coffee trees in the past 12 months, and zero otherwise. We also measure the number of coffee trees stumped based on farmers self-report to understand stumping intensity.

Moving along with the theory of change, if farmers adopt good agronomic practices, we should expect to see improvements in coffee and maize yields. We measure yields by administering plot and harvest modules that permit us to carefully measure the land size allocated to each of these crops as well as their harvested quantities.

Besides adoption of good agricultural practices and yield, asset holdings and household income are HereWeGrow’s key strategic indicators for assessing project impact. We measure asset holdings using an asset index, which aggregates the durable and productive assets owned by the household using a principal components analysis method (Sahn and Stifel 2003). Total household incomes are measured using detailed food and non-food consumption modules (because incomes are notoriously difficult to measure in contexts like these).³ We follow the guidelines in Deaton and Grosh (2000), by fielding detailed survey modules capturing households’ food and non-food consumption over a pre-specified period. These data are then aggregated into a single indicator capturing annual household per capita consumption-expenditures following standard approach outlined in Deaton and Zaidi (2002). Household consumption expenditures are reported on a per capita basis and in prices prevailing during the baseline interview. Due to the log-normal distribution (i.e., right-skewed) of consumption data, we will log this variable in the analysis.

Table 1. Primary outcome indicators

Variable	Variable type
Knowledge of good agricultural practices	Continuous
Adoption of good agronomic practices	Continuous (count)/Binary
Yield (coffee and maize)	Continuous, logged
Asset index (durable and productive assets)	Continuous
Household per capita consumption expenditure	Continuous, logged

We will also consider the direct measure of household income as a secondary indicator. The direct income measure is based on detailed gross income modules in the survey instrument: income from coffee production; income from maize production, income from other agricultural

³ A 2011 World Bank study from Uganda found that high-frequency farmer diaries resulted in average food production estimates that were 65% larger than those of the standard recall method (Deininger et al. 2012). Another World Bank study compared income and consumption reports from 20 developing countries and found that reported incomes were systematically lower compared to reported consumptions. Further analysis of the data suggested that income underreporting was most serious in agricultural income (Carletto, Tiberti, and Zezza 2022). It is partly for these reasons that virtually all national poverty statistics in low-income countries are based on consumption expenditure – which is considerably easier to measure accurately than income (Deaton and Zaidi 2002).

activities; wage income; non-agricultural business income, transfers and remittances. Household incomes are reported on a per capita basis and in prices prevailing during the baseline interview. Incomes are reported for activities managed by any member of the household as well as for activities managed by women only. As incomes are typically log-normally distributed, we log the household income variable. However, for income from female-led activities we consider a non-logged variable because this variable will likely include zeroes (logarithm of zero is undefined).

The secondary outcome indicators include alternative welfare indicators (household per capita income, food security) and access to inputs, technologies, and credit (Table 2).

Table 2. Secondary outcome indicators

Variable	Variable type
Household per capita incomes	Continuous, logged
Food Insecurity Experience Scale (FIES)	Continuous
Household Diet Diversity Score (HDDS)	Continuous (count)
Access to improved inputs and technologies (i.e., improved seedlings, improved seeds, fertilizers, pesticides)	Binary
Access to credit	Binary

3.4. Sample and statistical power

The sampling frame in this study is formed of coffee farming households residing in all eligible kebeles in the four woredas included in the study. We listed all eligible households, digitized these lists and randomly selected four households from each development group to be included in the evaluation sample. To balance the distribution of study arms within kebeles, we applied a stratified randomization approach defining kebele as the strata. Randomization was conducted by the research team using Stata before the baseline survey. Appendix A provides a detailed overview of the randomization and sampling approach.

The sample size was determined based on careful power calculations. We conducted power calculations using the Stata command `power twomeans`, setting the significance level at 5 percent and power at 80 percent, and allowing for 10 percent attrition between baseline and endline surveys. Power calculations are reported only for the primary outcomes. Detailed calculations together with the data sources are provided in the pre-analysis plan. Our estimates of standard deviations adjust for the reduced variance in outcomes when controlling for strata (kebele) fixed effects as we do in the final impact estimation.

We estimate statistical power corresponding to the comparison of each treatment arm to the control arm. Our study is adequately powered to detect improvements in the knowledge of good agricultural practices and subsequent adoption of them. The evaluation can detect a 0.34-unit improvement in knowledge of these practices (relative to the assumed baseline mean, this is 7 %). While we lack previous data to make assumptions about baseline levels of a composite index capturing the adoption of good agronomic practices, our study is powered to detect a 3-percentage point increase in coffee tree stumping adoption rate. Effects smaller than this are

unlikely to lead to meaningful positive effects on downstream indicators of yields and eventually to household welfare outcomes (e.g., per capita consumption).

Turning to these downstream effects, the MDEs for coffee yield is 29 percent, and 13 percent for maize yield. It is 13 percent for per capita consumption. While the MDE for coffee yield is on the high side and primarily driven by the relatively high ICC estimate (0.30), it is attainable if the intervention leads to a widespread adoption of stumping among treated farmers. Recent estimates from a similar setting in Ethiopia indicate that stumped coffee trees yield 2.8 kg/tree on average while the average yield from unstumped coffee tree is 2 kg/tree, suggesting a 40 percent yield impact (Hoppenbrouwers and van Haften, 2023).

These power calculations assume perfect compliance. Based on previous work evaluating digital extension interventions implemented by Digital Green in Ethiopia where it was found that videos were screened in more than 80 percent of the clusters (Abate et al. 2023), we expect a high degree of compliance. Assuming this level of compliance would inflate the estimated MDEs by 25 percent. For example, for (log) consumption, an 80 percent compliance would result in an MDE of 16 percent instead of 13 percent. Finally, we note that the final impact estimates will be based on an ANCOVA method, which will bring in additional power gains, particularly for outcomes displaying low autocorrelation such as per capital consumption and yields (McKenzie 2012).

3.5. Data and timeline

The quantitative data is the main source of information to measure the intervention's effectiveness and impact on indicators that can be statistically associated with the project.⁴ The quantitative data is collected using farm household surveys based on structured survey instruments developed by IFPRI in consultation with HWG and Digital Green. At baseline, the quantitative survey was designed based on the final log frame of the project and follow-up surveys will be conducted at midline (over the phone 12 months after the project launch) and at the endline (at the end of the intervention: four years after the start of the project). In these follow-up surveys, all households interviewed at the baseline will be visited. The midline questionnaire will be shorter in length and will focus on measuring project implementation performance (e.g., access to digital extension advisory, access and use of inputs and input credits) and adoption of good agricultural practices promoted by the project. Table 3 summarizes the timeline, particularly for the three rounds of quantitative data collection.

⁴ The evaluation team will also collect qualitative data from coffee and maize farmers and key stakeholders at the local level through focus group discussions (FGDs) and key informant's interviews (KIIs). The qualitative data will be mainly used to complement the quantitative data, as it provides more nuanced understanding of many of the outcome and impact indicators, narrating the evolution of changes in the words of farm households and local stakeholders.

Table 3. Timing of randomization, surveys, and intervention

Activity	Month, Year (tentative)
Randomization of clusters into study arms	November, 2023
Baseline survey	December 2023 – January 2024
Intervention begins	January, 2024
Midline survey	December 2024 – January 2025
Endline survey	December 2026 – January 2027

Note: The midline and endline timelines are tentative and may be pushed to April – May to collect information on stumping after the rejuvenation season is completed.

Baseline survey

As indicated in Table 3 above, the baseline survey was conducted in December 2023 and January 2024, after we obtained ethical approval from the International Food Policy Research Institute – Institutional Review Board (IRB #00007490). The fieldwork for the baseline survey was conducted by Greenpact under the supervision of Dr. Solomon Bizuayehu, working in close collaboration with the IFPRI team. Greenpact provided all equipment (data collection was performed on Android Samsung tablets) and led the organization of transportation and logistical services.

In the survey preparation phase, Greenpact staff translated the survey questionnaire into both Amharic and Afaan Oromo, coded in SurveyCTO to be conducted as a computer assisted personalized interview (CAPI), and led the testing and refinement of the program. The activity manager, the quality control specialist, and the field coordinators also revised the translation and phrasing of the questionnaires and supporting implementation materials from English into the local languages based on the feedback received during training and piloting. Greenpact also led the process of obtaining appropriate approvals for survey work from local stakeholders. Greenpact hired highly experienced, knowledgeable, socially accepted, and well-connected guides in each target sample kebeles to maximize the participation rate of households.

The IFPRI team conducted the enumerators training, which took place between November 25 and December 01. The training ensured that all enumerators interpreted the survey questions consistently and asked the questions in the prescribed manner. Greenpact’s survey coordinator and field supervisors were trained to fully explain the purposes and importance of the baseline household survey in simple terms and to reassure the respondent about the confidentiality of information. Quality control specialists and database managers were trained to run high-frequency consistency checks on the data from the office using Stata and conduct audio checks for sample interviews, throughout the data collection process.

The training process consisted of a combination of classroom training, mock interviews, pilot interviews, and debriefing. During the training, the IFPRI team explained the meaning and content of each question. They defined the code of conduct, the composition, and organization of the team, and the communication channels to be employed within the team. There was also a two-day CAPI-based mock interview session for all the survey team members at the end of the

conceptual training session. Before launching the main data collection, the survey team then conducted pilot interviews in a sample of households outside the main evaluation sample, in nearby kebeles that are not part of the study. During the pilot survey, Greenpact interviewed about 78 households. The objective of the pilot was to test the programmed version of the survey instrument and glean feedback on the length of the interview, the phrasing, context, and framing of the questionnaire, the quality of the translation, and strategies for approaching rural households. Pilot testing also provided the team with an opportunity to practice appropriate interviewing behavior, appropriate use of the CAPI program and tablets, and participation in the team's field routines (review of the data by supervisors, completion of field control sheets by supervisors, and distribution of work assignments and coordination by supervisors).

During the baseline survey, Greenpact regularly updated the IFPRI impact evaluation team regarding data collection activities and reported the progress of the data collection and challenges encountered on a regular basis. To aid in real-time data visualization, the IFPRI team employed a data monitoring dashboard embedded in Google Sheets. Several additional strategies were also employed to ensure the collection of high-quality data: random set of call-backs or visits by the survey coordinator, high frequency check using Stata do-files, and audio checks. The use of electronic data collection (CAPI) also provided additional opportunities for quality control. During the design stage, IFPRI's research officer in collaboration with the programmer exploited the capabilities of the SurveyCTO program to incorporate automatic skip patterns and constrain responses so that the enumerators were not able to continue recording responses when the data was clearly incorrect, invalid, or inconsistent.

The baseline survey targeted 3,000 households (i.e., 4 households in each 125 development groups), with the survey team successfully completing interviews at 2,995 of them. Five households were excluded from the evaluation sample because they did not produce coffee and thus were ineligible.⁵

⁵ There are also three households for which the household roster information is incomplete. We have included these three households in the sample but omit them when we report on household demographics or outcomes, such as consumption or poverty that are expressed in per capita terms.

4. Results and discussions

4.1. Household demographics and socio-economic characteristics

Table B1 in appendix B reports the summary statistics of key household characteristics across the six study arms. We report the pooled mean, the mean in each treatment arm and the corresponding standard errors (clustered at the development group level). Given the randomized design, we expect observable and unobservable characteristics to be balanced across study arms. We do not observe large differences in the means across the study arms, indicating reasonable balance across the study arms.

Table B2 reports more formal balance tests: F-tests for joint significance across all arms for each variable. These tests control for strata fixed effects. For study arms 1 to 5 these fixed effects are constructed at the kebele level (Column 1). Clusters forming the study arm 6 (pure control) originate from different kebeles than the clusters in other study arms. When selecting the clusters for pure control we ensured balance across woredas. Therefore, to assess the balance between study arm 6 and the other arms, we control for woreda fixed effects (Column 2). Across the 240 statistical tests reported in Table B2, only 14 appear statistically significant at the 5-% level. Given the successful randomization achieving balanced characteristics across groups, and since the interventions had not begun at the time of the baseline survey, we will characterize the households using the pooled sample in this baseline report.

Household demographics

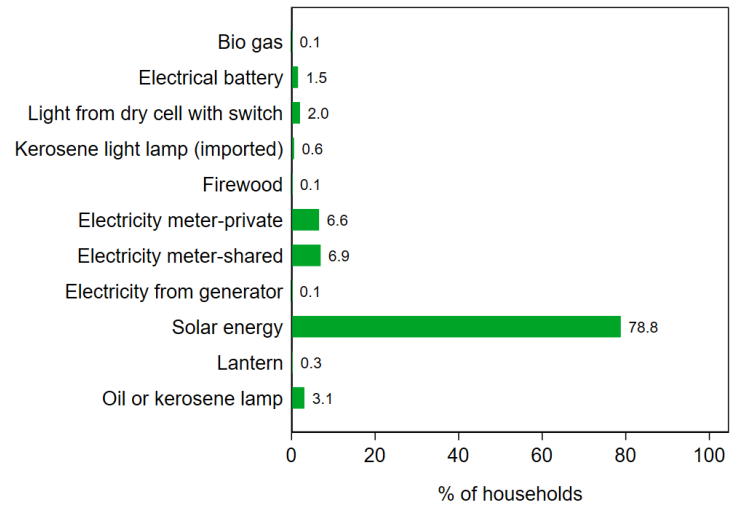
The mean household size in our sample is 5.5 members and median 5 members. However, these averages mask considerable heterogeneity. The smallest household only has one member while the largest has 18 members. Less than nine percent of the households are female-headed. The mean age of the household head is 45 years and 43 percent of the household heads are illiterate (cannot read or write).

Housing characteristics and remoteness

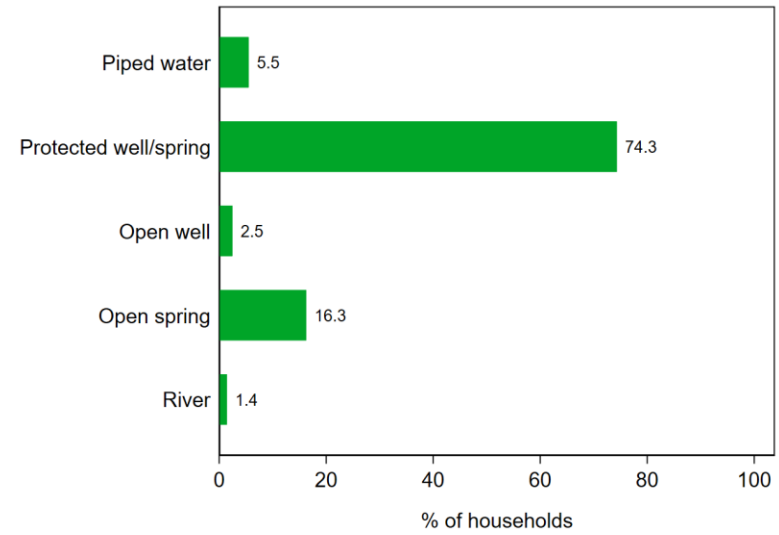
Access to electricity from the main grid is rare and most households (79%) use solar energy to light their houses after sunset (Figure 2A). Nearly 74 percent of the households obtained water from protected well or spring and 16 percent from an open spring. Access to piped water is rare, reported only by 5.5 percent of the households (Figure 2B). Ninety-five percent of the households had iron corrugated roofs, others mainly thatched roofs (Figure 2C). As for sanitation, pit latrines are common, often without a slab (Figure 2D), though about three percent of the households defecated in the open, in nearby fields or forests. Virtually all households (99.2 percent) reported firewood (collected by themselves) as their main source of cooking fuel (no graph reported).

Figure 2. Housing characteristics

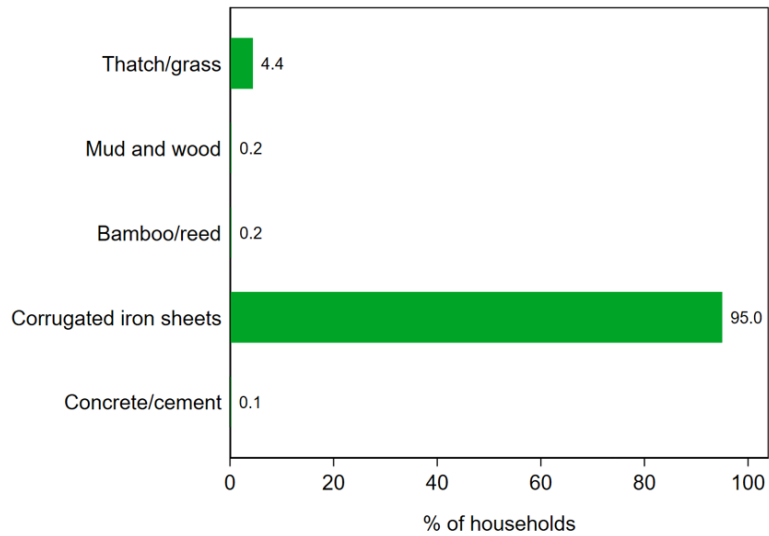
A. Source of light



B. Water source



C. Material of the roof



D. Toilet type

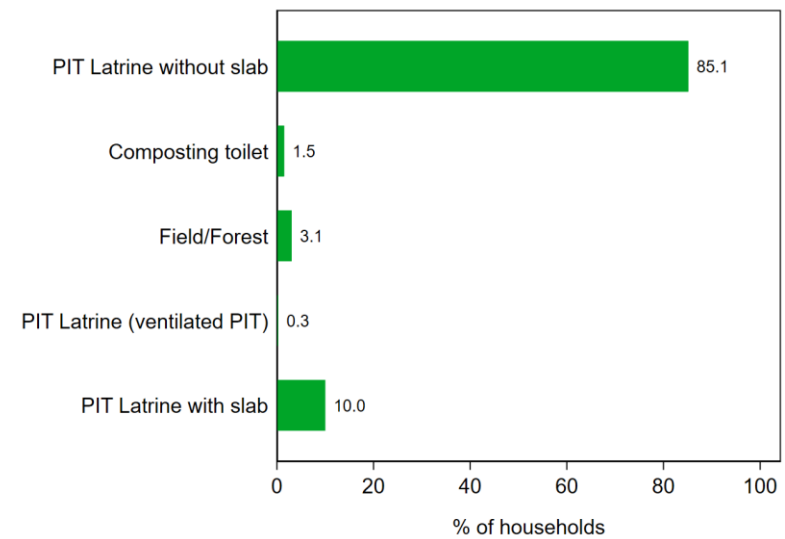
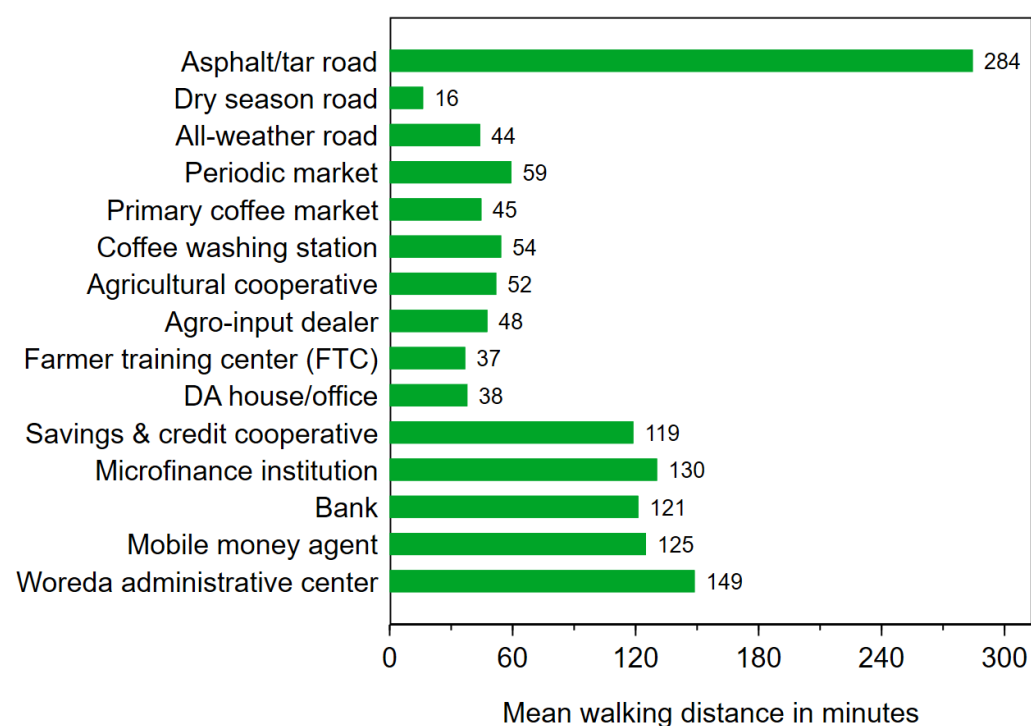


Figure 3 reports the mean walking distances in minutes to roads, agriculture offices and finance institutions. The average household in our sample needs to walk more than 4.5 hours (284 minutes to the nearest asphalt (or tar) road. Other types of roads are closer with a road accessible during the dry season at 15-minute walking distance on average and an all-weather road at 45-minute walking distance. The nearest periodic (typically weekly) market and the primary coffee market are both less than 1 hour walk away from household’s home location, on average. The same is true for coffee washing stations, agricultural cooperatives, and agro-input dealers. Farmer training centers and DA office are both relatively accessible with the average household reporting a half-hour walking distance to these places. The average distances to various financial institutions (savings and credit cooperatives, MFIs, Banks, mobile money agents) is around two hours. The average walking distance to the woreda administrative center is 150 minutes, i.e., more than two hours.

Figure 3. Average walking distances to roads, markets, agricultural offices and financial institutions



Durable and farm assets

Figure 4A shows the share of households owning different types of durable assets. More than 80 percent of the households report owning solar panel or battery and a phone (most often a mobile). Many households report owning various types of furniture. Very few households own a bicycle but nearly eight percent report owning a motorbike. Seven percent have a TV, typically connected to a satellite dish.

Figure 4. Ownership of durable assets

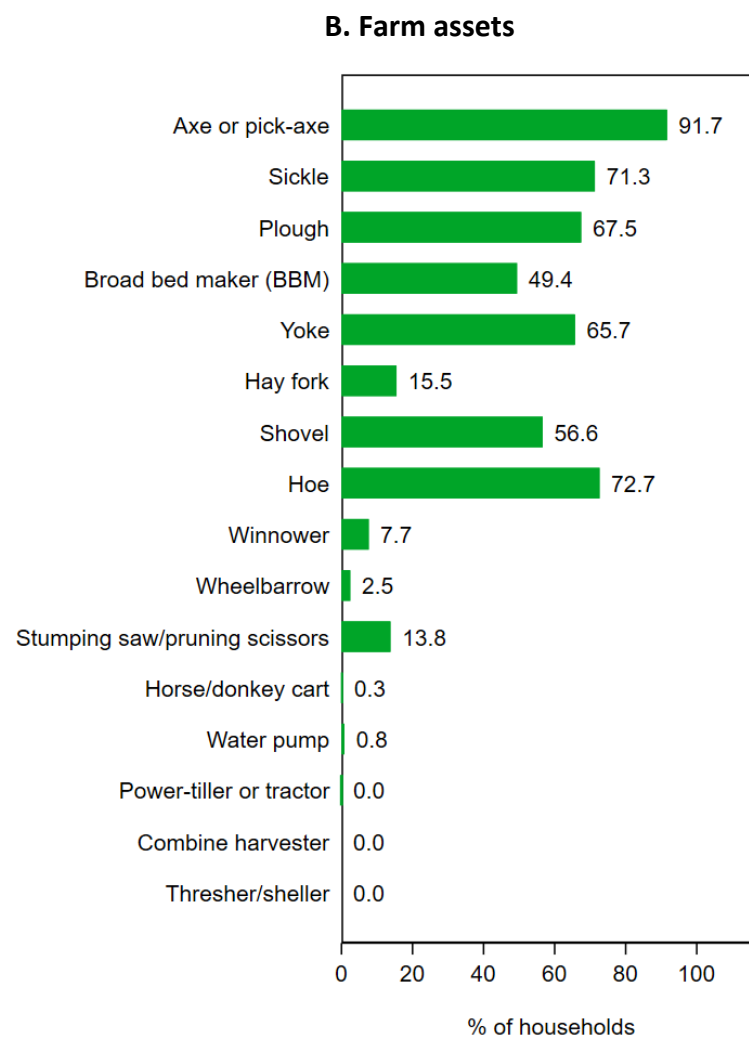
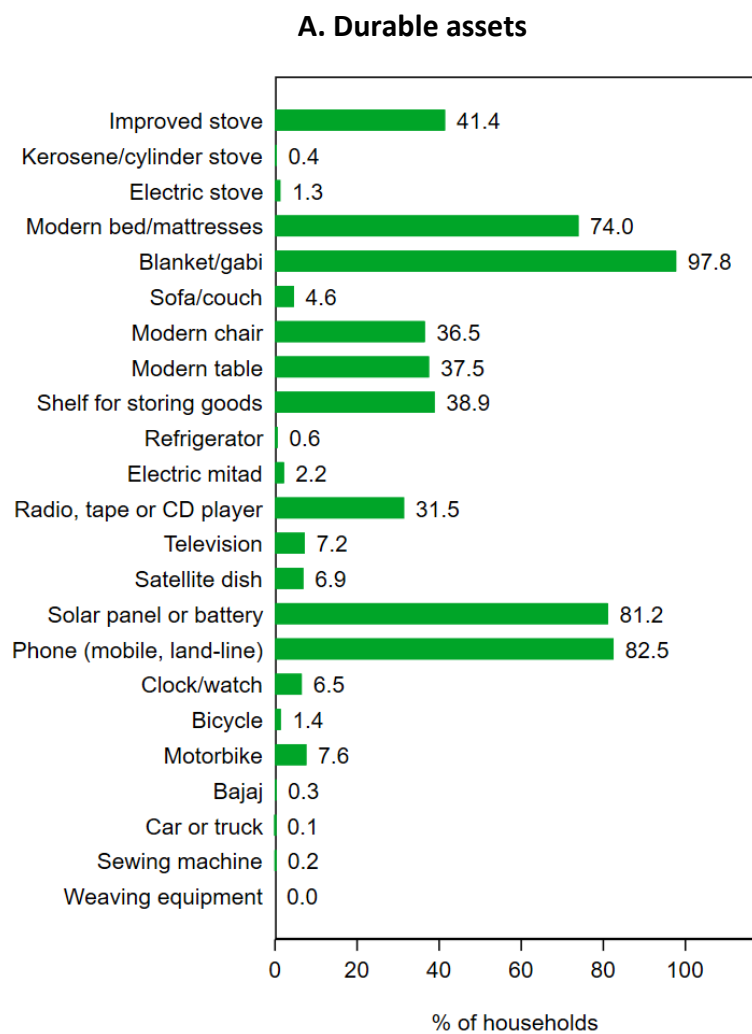
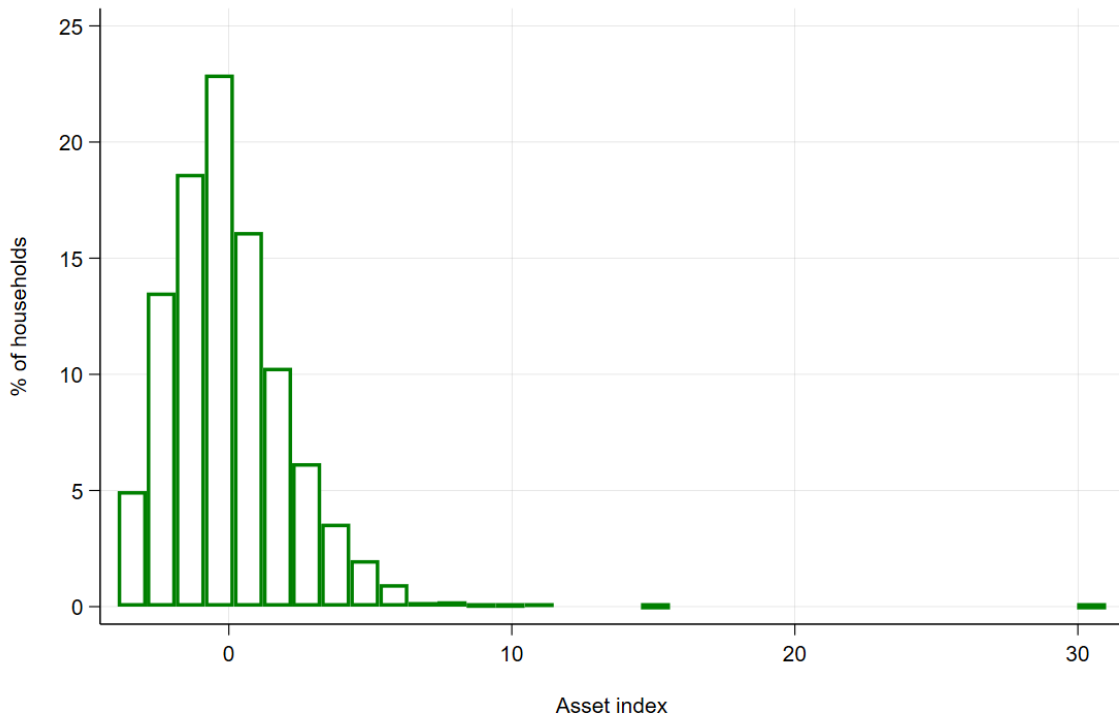


Figure 4B above reports on the ownership of various farm assets. More than 90 percent of the households in our sample have an axe, more than 80 percent own a sickle and 67.5 percent own a plough. More than half of the household report owning a broad bed maker (BBM), yoke, shovel and hoe. Ownership of other types of farm assets is rare.

We aggregated the ownership of all durable and farm assets into single value using a principal components analysis method (Sahn and Stifel, 2003).⁶ By construction, the sample level mean value of the resulting index is zero. The magnitudes of the index represent the relative positioning of households in terms of asset-based welfare. However, the specific numerical values of the index are abstract and do not have a direct, standalone meaning beyond this relative positioning. With these caveats in mind, Figure 5 shows the full distribution of the asset index. The index values are tightly clustered between -5 and 5, although a small number of households have values larger than 10. At Endline, we will use the asset index to assess the degree to which asset-based welfare varies across study arms.

Figure 5. Distribution of the asset index values



Note: N = 2,995 households

⁶ We used the number of individual assets owned by the households. Before aggregating using the PCA method, we winsorized all individual asset components at the 99-percentile.

Agricultural land

Table 4 provides the mean and median land sizes operated by the household. The mean (median) operated agricultural land area in our sample is approximately 0.97 ha (0.75 ha), out of which 0.51 ha (0.31 ha) is allocated to coffee farming, 0.27 ha (0.25 ha) to maize farming and 0.19 ha (0.06 ha) to other crops.

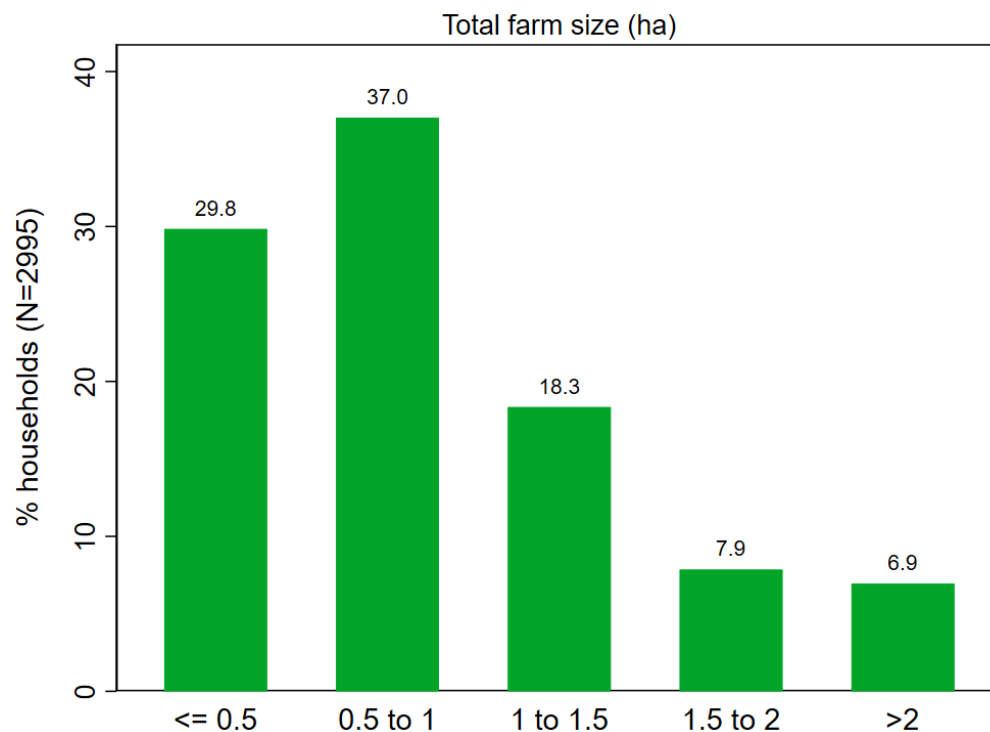
Table 4. Farm area, by study arm

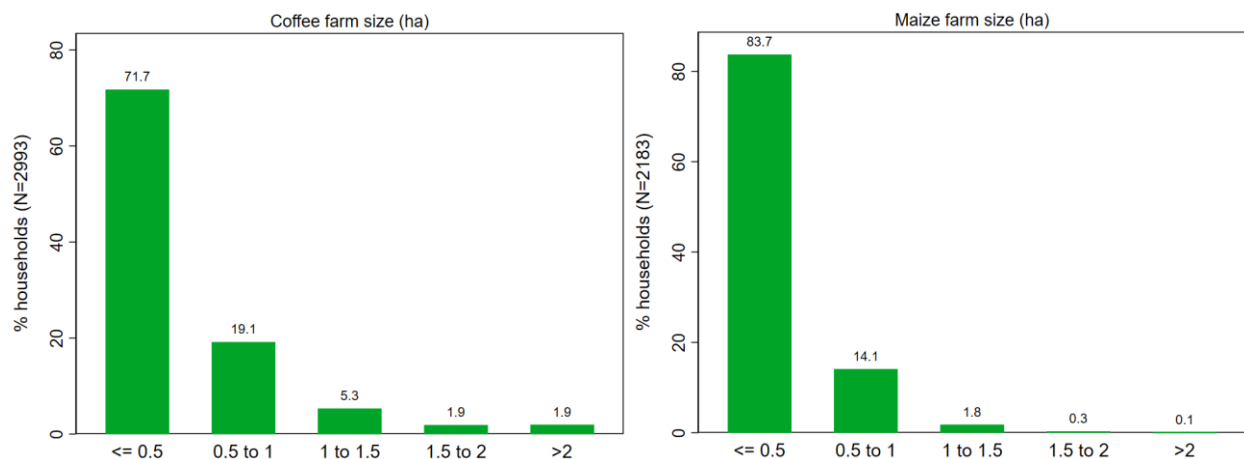
	Mean	Median
Total farm area (ha)	0.97	0.75
Coffee farm area (ha)	0.51	0.31
Maize farm area (ha)	0.27	0.25
Other crops farm area (ha)	0.19	0.06

Note: 2,995 households.

Figure 6 categorizes total land holdings into 0.5 ha bins and shows the share of households in each bin. This analysis reveals that about 25 percent of the households operate less than 0.5 ha whereas about 17 percent of the households operate more than 1.5 ha of land. The results for coffee and maize plots indicate that most households manage plots that are less than 0.5 ha (72 percent of households in the case of coffee plots and 84 percent of households in the case of maize plots).

Figure 6. Farm size categories, total and coffee and maize plots





Note: N = 2,995 households and ha refers to hectare.

4.2. Coffee and maize production, sales, knowledge and adoption of recommended practices

Coffee production and sales

Table 5 reports on coffee production, sales, and revenue during the 2023 (2015/16 E.C.) season. The baseline questionnaire asked respondents to estimate the amount of red cherry and jenfel they had harvested from each plot during the 2023 (2015/16 E.C.) production season. We converted the reported quantities of red cherry and jenfel to green beans using the following conversion factors from the Coffee Atlas of Ethiopia (Davis, Kebebew Challa, Williams, 2018, uchicago.edu): 2 kg of dried cherry coffee (jenfel) equals 1 kg of clean (green bean) coffee, and 5 kg of fresh cherry coffee equals 1 kg of clean (green bean) coffee. Aggregating the quantities across plots, the average farmer in our sample produced 157 kg (median: 86 kg) of green beans during the 2023 season. The mean total production of cherry was 246 kg (median: 100 kg), and for jenfel, it was 216 kg (median: 100 kg). The green bean yield (total production amount per hectare) was 348 kg/ha (median: 262 kg/ha). This is equivalent to the national average coffee yield, which was 687 kg/ha dried whole cherries (343 kg/ha of green beans) during the 2021/22 meher production season. At the woreda level, Shabe Sombo stands out as the most productive area, with an average yield of 584 kg of green beans per hectare. This yield is more than double that of Mana and Limu Kosa and about 58% higher than that of Seka Chekorsa (Table C5 in Appendix C).

We also asked households to report the amount of coffee they had sold by the time of the interview. Cherry is typically sold immediately after harvest (October-December), so most of these sales are captured in our data. However, jenfel is sold throughout the year and, as such, serves as an important form of saving in this context. With these considerations in mind, by the time of the interview, the mean amount of cherry coffee sold by the farmers was 160 kg (median: 70 kg), and jenfel coffee was 66 kg (median: 0 kg). These sales translate to mean revenue of 4,689 birr from selling cherries and 4,622 birr from selling jenfel, totaling 9,311 birr. At the time of the survey, cherry beans were sold at 29 birr/kg (median: 30 birr/kg), while the price of jenfel was

around 70 birr/kg (median: 65 birr/kg). While sizable share of red cherries are sold at roadside (38 percent), kebele market (27 percent), and wet mills (19 percent), most of the jenfel is sold at the woreda/district (51 percent) and kebele (31 percent) markets (see the place of sales by coffee type in Figure F1 and F2 in the annex).

Table 5. Household level coffee production, sales, and coffee revenue during the 2023 (2015/16 E.C.) production season

	N	Mean	Median
<i>Coffee production (kg), conditional on producing</i>			
Total cherry production	2995	246	100
Cherry per hectare	2993	528	340
Total jenfel production	2995	216	100
Jenfel per hectare	2993	482	345
Total green bean coffee production	2995	157	86
Green bean coffee per hectare	2993	348	262
<i>Coffee sales (kg)</i>			
Share sold cherry (%)	2995	68	NA
Share sold jenfel (%)	2995	35	NA
Cherry sold (kg)	2995	160	70
Jenfel sold (kg)	2995	66	0
<i>Coffee revenue (in birr), conditional on selling</i>			
Cherry revenue	2995	4,689	2,000
Jenfel revenue	2995	4,622	0
Coffee revenue per hectare	2993	21,197	12,000
Cherry revenue per hectare	2993	9,806	6,000
Jenfel revenue per hectare	2993	11,018	0
Total coffee revenue	2995	9,311	3,960
<i>Coffee prices</i>			
Cherry price (birr/kg)	2049	29	30
Jenfel price (birr/kg)	1036	70	65
Green beans (equivalent) price	2420	143	148

Note: The disaggregation of cherry and jenfel production aims to highlight their relative importance. However, the per-hectare cherry and jenfel productions do not reflect yield accurately, as a farmer may harvest cherries and sell their produce immediately as red cherries and/or process part of the harvest from the same plot to jenfel. Therefore, to assess yield, it is essential to refer to the quantity of green coffee beans (clean coffee) per hectare.

Knowledge of coffee best practices

The baseline survey included a quiz testing farmers' knowledge of recommended coffee farming practices. Table 6 shows the percentage of households responding correctly to these knowledge questions in each theme. Farmers showed a high degree of understanding of recommended practices for shade management and stumping methods. In contrast, only 38.5 percent of farmers know that weeding by digging under tree canopy was not a recommended practice. Out of a maximum of seven points on the knowledge quiz, the average farmer in our sample scored

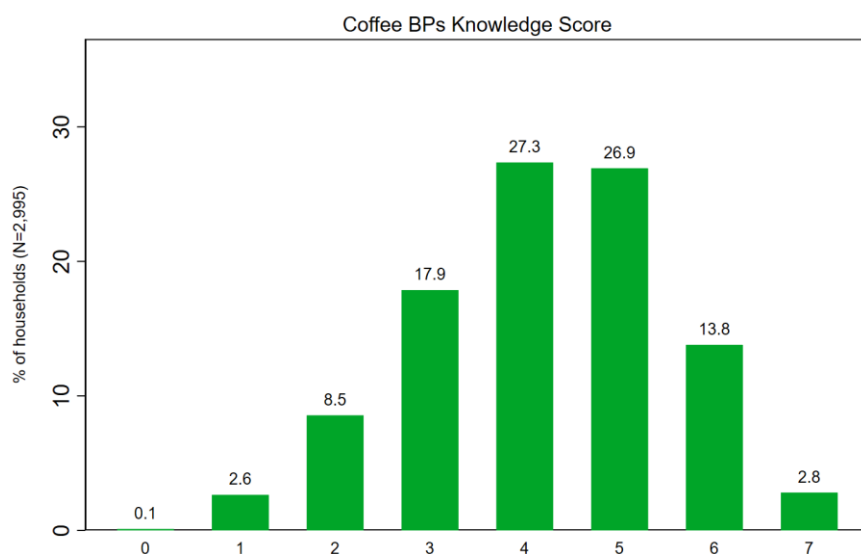
4.2 points. Figure 7 shows the full distribution of knowledge scores: 71 percent of the households scored at least four out of the maximum seven.

Table 6. Knowledge of recommended coffee farming practices

Recommended practice	Description	% of households responding correctly
Weeding	Not digging under tree canopy	38.5
Sucker selection	Suckers left at first & second thinning	43.4
IDM	At least one disease control method	68.5
Nutrition	Feeding coffee trees	69.5
Stumping method	Cutting method during stumping	76.1
Shade management	Shade cover knowledge	94.3
Planting coffee seedling	Seedling planting hole preparation time	29.3

Note: IDM stands for Integrated Disease Management.

Figure 7. Recommended coffee farming practices: distribution of knowledge score



Adoption of coffee best practices

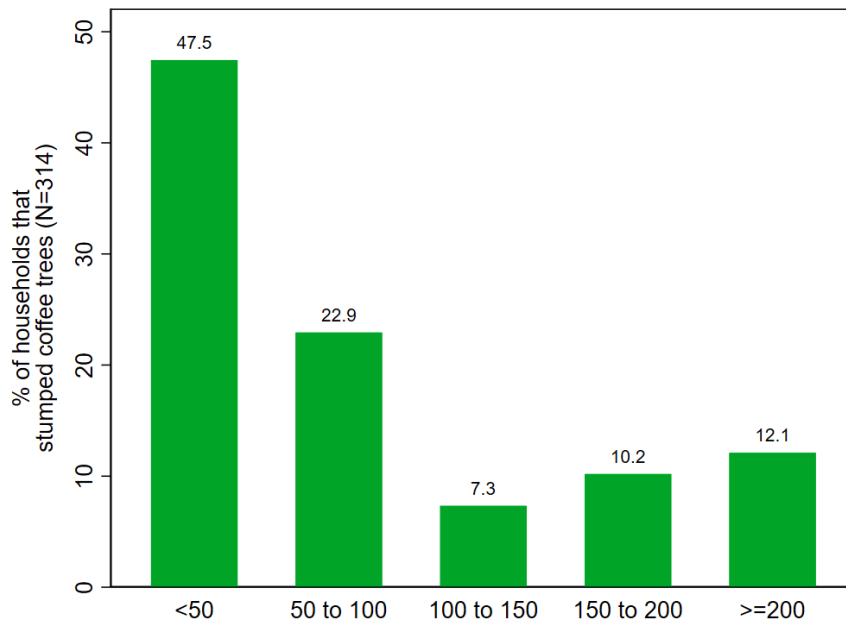
The ICoFs project is primarily designed to promote best agronomic practices on coffee plots and this subsection presents the results on the adoption rates of these key agronomic practices (i.e., rejuvenation/stumping, planting new trees, coffee nutrition, weeding, integrated disease management, and shade trees management).

Rejuvenation: rejuvenation or stumping is a key best practice that entails properly cutting old and unproductive coffee trees at its base for a complete renewal. We measured coffee stumping based on farmers self-report and visual observation by enumerators on the reference/visited

coffee plot. The baseline adoption measure considered stumping of coffee trees by households over a three-year period prior to the start of the project, given that coffee farmers in Ethiopia have rarely been taught on how to stump their coffee trees. With this measure, about 10 percent of the households had stumped their coffee trees in the three stumping seasons/years prior to the start of the project (Figure 14). While there is no much difference on stumping adoption rate across the study arms, we observed significant variation in stumping adoption rate across woredas. For instance, 20 percent of sample households in Mana woreda adopted stumping at baseline, which is significantly higher than the 6-7 percent adoption rate in the other three target woredas.

Looking at intensity of stumping, estimated number of coffee trees by small number of households (314) that had adopted stumping range from 1 to 1000 coffee trees, with a mean of 103 trees and a median of 57 trees. Figure 8 presents stumping intensity by category and about half stumped less than 50 coffee trees, while only a third stumped more than 100 coffee trees.

Figure 8. Stumping intensity

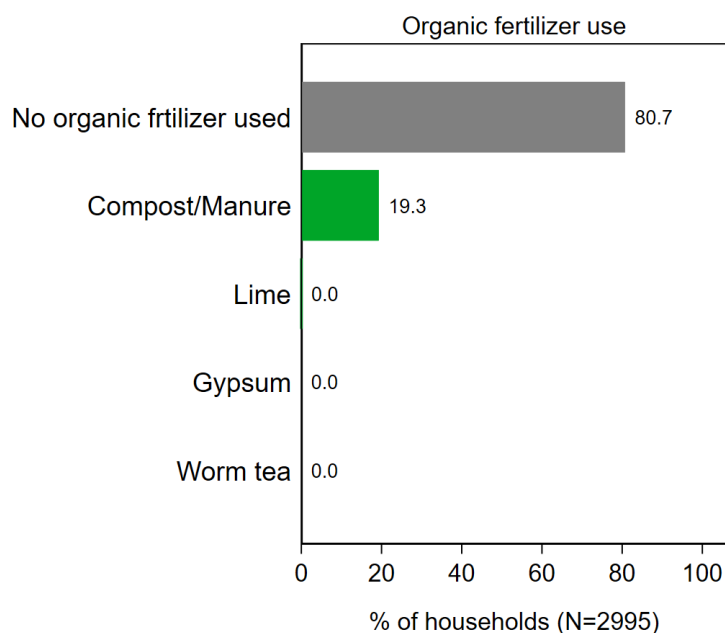


Planting coffee seedling: Besides rejuvenation, farmers also plant new seedlings to renew or expand their coffee farms. The results from the baseline survey show that more than half of the households (53 percent) planted new coffee seedlings in the 12 months prior to the survey. The number of new coffee seedlings planted by an average household is estimated at 120 and most of these seedlings (66 percent) were sourced from their own farm. A sizable share of households also source seeding from private (20 percent) and government (10 percent) nurseries (see Table C7 in the annex). In addition to planting new coffee seedlings, 28 percent of households reported having a coffee nursery (seed bed) in the past 12 months. A household is considered an adopter if it either planted new coffee trees or owns a coffee nursery within a 12-month period. Using this criterion, 51 percent of households are qualified as adopters.

Organic fertilizer (coffee nutrition): coffee nutrition is another key best practice that the project promotes to increase coffee productivity. The nutritional status of coffee plots was measured by the use of organic inputs such as compost and manure on the reference/visited plot. A household is considered as an adopter of coffee nutrition practice if the household had applied at least one organic soil input during the recent production session prior to the baseline survey. Given that compost production and use is seasonal, the use of compost or manure was verified by looking at the compost heap, pit, or pile of compost as evidence of composting.

With this measure, only 19.3 percent of visited plots satisfied/passed the coffee nutrition practices adoption requirements (Figure 14). The result indicates that there was a limited use of organic fertilizers among farmers (Figure 9).

Figure 9. Organic fertilizer uses

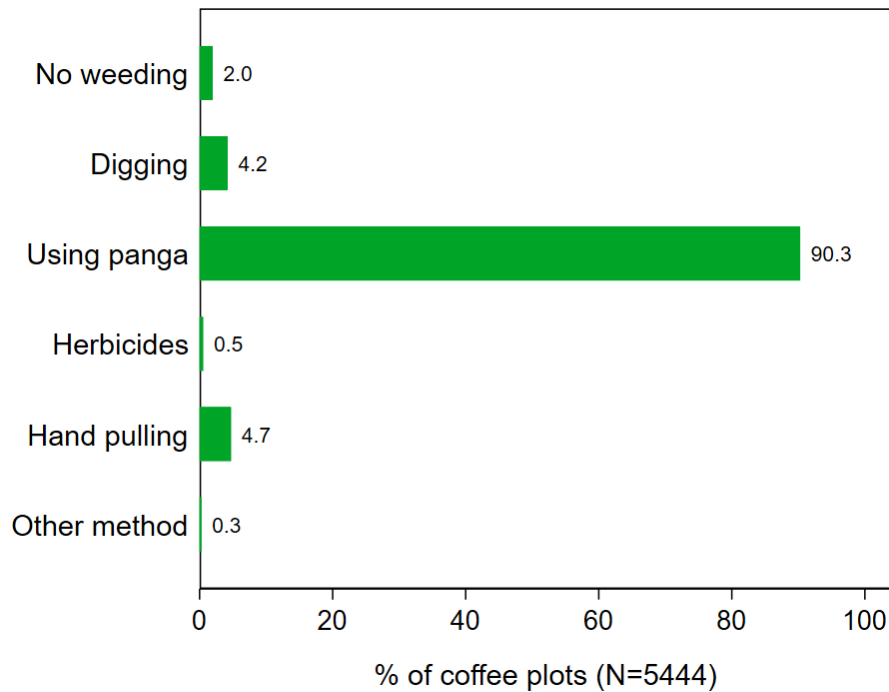


Weeding: The project also promotes proactive weed control as a key best practice to keep coffee plots weed-free throughout the year to increase coffee tree yields. If allowed to grow in coffee areas, weeds can seriously affect yields, since weeds aggressively compete with coffee trees for soil moisture, nutrients, and sunlight. The adoption measure for weeding best practice considers several factors: weeding frequency, presence of weeds, heights, and weeding methods (details of adoption rules found at Appendix G). A farmer is considered as an adopter, if the visited or reference plot is weeded at least twice a year; there are few or no weeds under the tree canopy; any existing weeds are less than 30cm tall, and the area under the tree canopy has not been dug to clear weeds.⁷ Based on this measure, more than half of the households (57%) had adopted the weeding best practice and there are no meaningful differences across study arms (Figure 14).

⁷ Digging under the tree canopy is not a recommended practice since it can damage the all-important feeder roots, introduce coffee wilt disease, and leave the soil open to erosion.

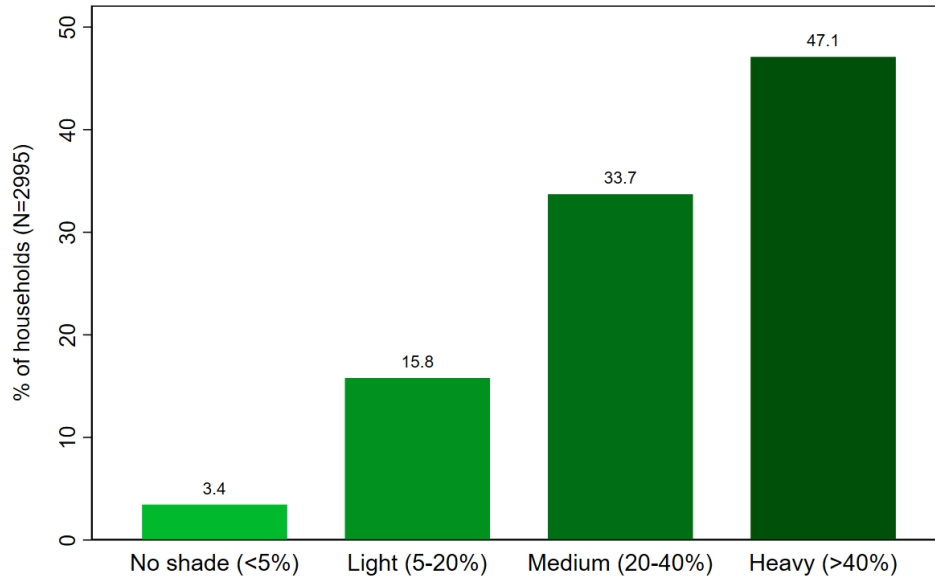
Like that of rejuvenation, adoption of weeding relatively higher in Mana woreda (68 percent) compared to the other project woredas (46 – 57 percent). Weeding with panga is the most common weeding method used in the target woredas (Figure 10).

Figure 10. Weeding methods



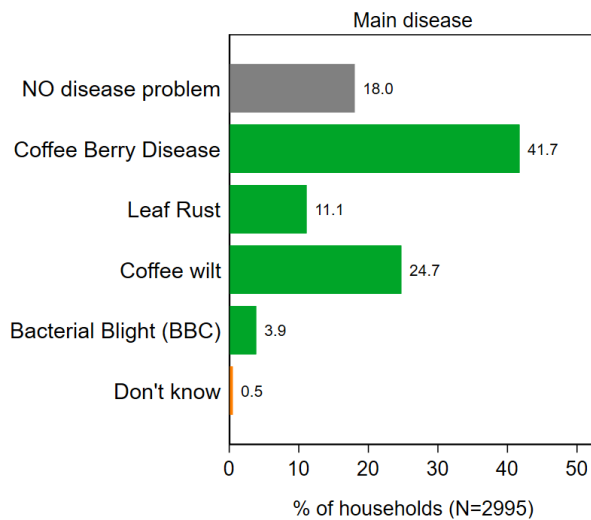
Shade management: The project promotes growing coffee with shade trees both for economic (increasing yield and quality of coffee beans) and environmental (maintaining ecosystem services) reasons. The adoption of shade trees is measured by the estimated share of coffee trees covered with shade on the visited coffee plot. Households are considered adopters if the estimated shade level is 20 percent or more. Based on this measure, the baseline results show a higher adoption rate of shade trees at 82 percent (Figure 14). However, only 34 percent of the coffee plots have a recommended shade level, which is 20-40 percent (see Figure 11). It is worth mentioning that baseline shade adoption rate is significantly lower in Shabe Sombo woreda (57 percent) compared to the other target woredas (85 – 90 percent).

Figure 11. Shade level



Integrated disease control method (IDM): Coffee diseases (e.g., coffee berry disease, coffee wilt disease, coffee leaf rust) can result in severe yield loss and quality deterioration if not properly controlled. The project promotes a wide range of methods farmers can use to prevent and control the incidence of diseases on their coffee farm. We first asked farmers about their main disease problems in the last 12 months. While 18 percent of the households reported not having disease problems, 42 percent and 25 percent of the households identified Coffee Berry Disease and Coffee Wilt, respectively, as the most prevalent disease problem on their coffee farm (Figure 13, right panel).

Figure 12. Main disease problems

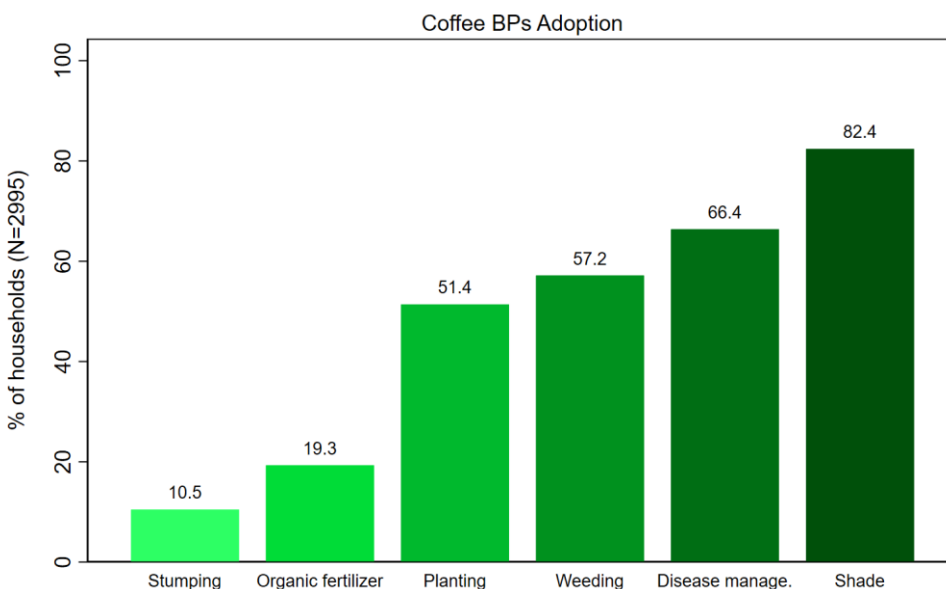


Farmers were also asked how often they inspect their coffee farms for diseases and pests. Approximately 37 percent reported inspecting weekly, 32 percent every two weeks, and the remainder inspect monthly or less frequently.

Households are considered adopters if they reported using at least one method to control disease when a disease was encountered or, in the absence of disease incidence, inspected their coffee farms weekly or more often. Accordingly, 66 percent of the sample households were adopters of integrated disease management (Figure 14). We observed that relatively large share of households (74 percent) in Mana woreda used at adopted integrated disease management compared to the other target woredas (11-18 percent). Uprooting infected coffee tree was the most common method of pest and disease control.

Figure 14 presents the summary of the key best practices adoption rates, and the results are mixed. More than 80 percent of the households reported using shades to protect their coffee trees from excessive heat, 66 percent practiced disease management and close to 60 percent reported weeding their coffee farms in the past 12 months. Meanwhile, more than 50 percent of the farmers planted new coffee trees. Other recommended best practices, such as soil fertility and plant nutrition management were followed/practiced by about 20 percent of the farmers. Only about 10 percent of the farmers reported to have stumped (rejuvenated) their coffee trees in the last three years (see Section 5.1.).

Figure 13. Summary of best practice adoption



Note: IDM stands for Integrated Disease Management. The rejuvenation result is based on stumping adoption over three seasons, 2021 – 2023.

Maize production and sales

More than 70 percent of the farmers in our sample produced maize (Table 7). The average maize farmer reported producing 694 kg of maize (median: 500 kg), and the average yield was about 2,000 kg per hectare (median: 1,714 kg/ha). (This is much lower than the national and zonal average yield, which is estimated at more than 4,000 kg per hectare). Most of this maize was for home consumption, as only 13 percent of the farmers reported selling or planning to sell the maize they had harvested. Among those who sold maize, the average quantity sold was 426 kg

(median: 300 kg), translating into 12,420 birr (median: 8,000 birr) in revenue per farmer. These farmers sold a kg of maize for 33 birr (median: 30 birr/kg).

Table 7. Household level maize production, sales, and maize revenue during the 2023 (2015/16 E.C.) production season

	N	Mean / %	Median
Share of household produced maize	2,995	73%	n/a
Total maize production (in kg)	2,183	694	500
Maize production per hectare (in kg)	2,181	2,016	1,714
Share of household sold (plan to sell) maize	2,995	13%	n/a
Quantity of maize sales (in kg)	399	426	300
Maize sales revenue (in birr)	399	12,420	8,000
Maize revenue per hectare (in birr)	399	22,753	16,000
Maize sales price (birr/kg)	366	33	30

Knowledge and adoption of recommended maize production practices

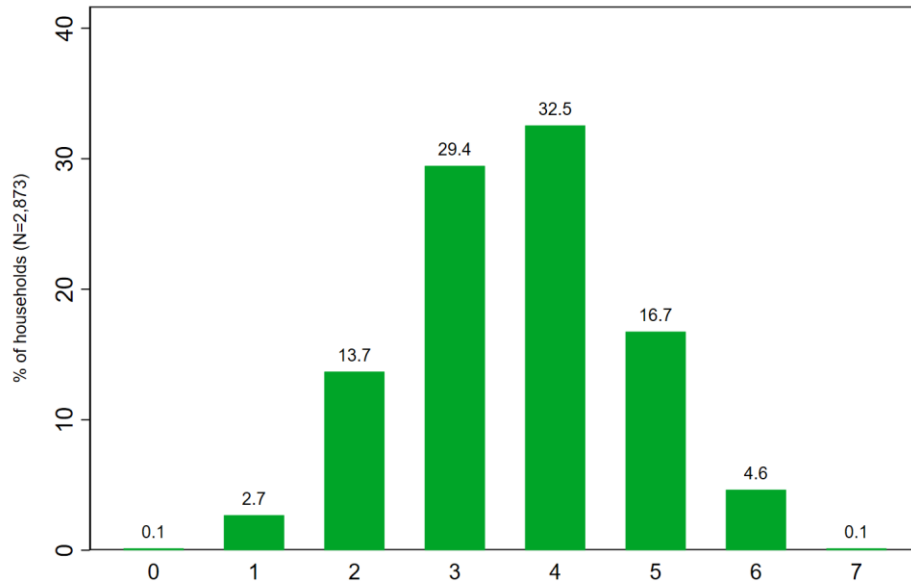
The baseline survey included a quiz testing farmers' knowledge of recommended maize farming practices, like the quiz administered on coffee farming. (However, due to a programming mistake in the questionnaire, this sub-module was not administered to the first 122 households interviewed in the survey). Table 8 shows the percentage of households responding correctly to these knowledge questions in each theme. Farmers showed a good grasp of recommended practices for ploughing, NPS/DAP fertilizer application and Fall Armyworm (FAW) control methods. However, their understanding of recommended practices for UREA application, inter-cultivation and row planting was generally limited. Out of a maximum of seven points on the knowledge quiz, the average farmer in our sample scored 3.6 points. Figure 15 shows the full distribution of knowledge scores: 54 percent of the households scored at least four out of the maximum seven.

Table 8. Knowledge of recommended maize farming practices

Recommended Practice	Description	% of households responding correctly
Ploughing	Ploughing knowledge (3 to 5 times)	95.8
Row planting	Row planting (row width, dist. between plants, seed depth and seed density)	38.1
NPS/DAP application	Application time and amount applied	75.6
UREA application	Application time and amount applied	5.5
Inter-cultivation ('shilshalo')	Inter-cultivation is not recommended	33.0
Fall Armyworm (FAW) control	Know at least one FAW control method	63.0
Storage	Know at least three recommended practices	50.1

Note: Due to a programming mistake, this data is missing for 122 households. N = 2,873 households.

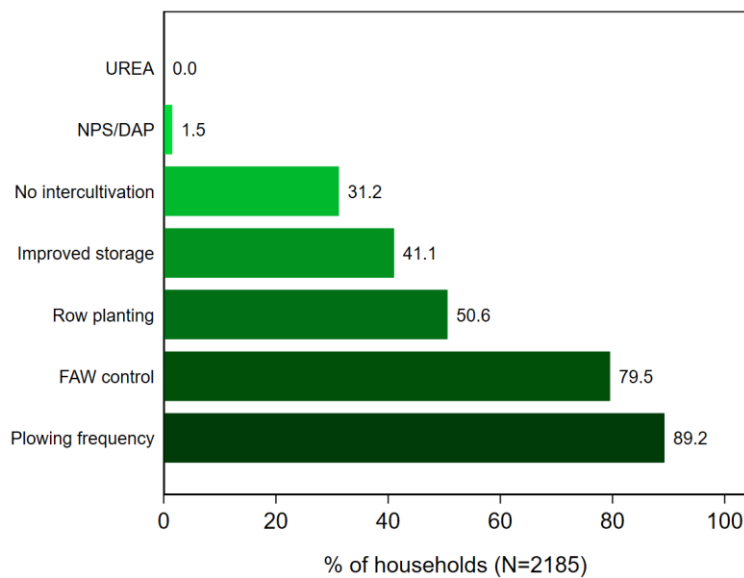
Figure 14. Recommended maize farming practices: distribution of the knowledge score



Note: Due to a programming mistake, these data are missing for 122 households. N = 2,873 households.

Figure 16 reports on the adoption of recommended maize farming practices. Around 80-90 percent of the farmers cultivating maize reported plowing more than three times before sowing, and applied FAW control method when there was incidence or frequently inspected their maize field. However, most farmers did not apply the recommended per-hectare amount of NPS, or urea. Typically, the applied amounts fell below the recommended amounts with the average farmer applying 33 kg/ha of NPS/DAP and 33 kg/ha of urea when the recommended amounts are 100-150 kg/ha and 150-250 kg/ha, respectively.

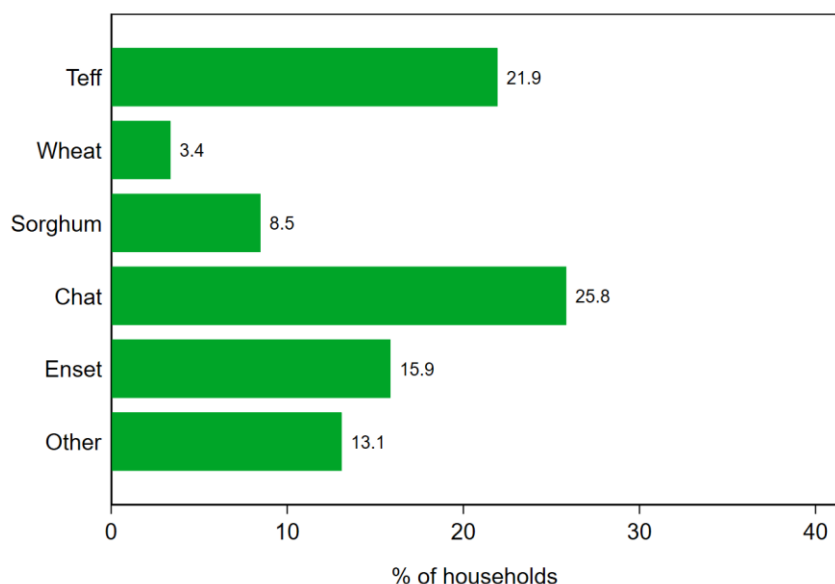
Figure 15. Adoption of recommended maize farming practices



4.3. Production of other crops

More than half (55%) of the households reported growing crops other than coffee and maize. Chat (khat) and teff are the most common other crops, cultivated by more than 20 percent of the farmers in our sample (Figure 17). These are followed by enset (16%), sorghum (8.5%) and wheat (3.4 %). Other individual crops were less common.

Figure 16. Other crops cultivated by the farmers



Note: N = 2,995 households.

4.4. Livestock ownership

Livestock in this context serves many functions. They act as a form of buffer stock savings that can be liquidated into cash during difficult times. Livestock products, such as dairy, eggs, and honey, generate cash income that can be used to purchase food, agricultural inputs, or school supplies. Finally, large livestock are used as draft power in crop agriculture.

Figure 18 shows the share of households owning different livestock species. At the time of the interview, 85 percent of the households reported owning livestock. More than 50 percent of the households owned oxen/bulls or cows and calves. Pack animals, such as horses, donkeys, and mules, were owned by 25 percent of the households, while more than 40 percent raised chickens. Goats and sheep were owned by 13 percent and 32 percent of the households, respectively. 22 percent of the households reported owning at least one beehive. No household reported owning camels, and only one household raised pigs. Figure 18 shows the median number owned, conditional on owning the livestock type. These vary from 1 ox/bull or pack animal to 3 chickens or 3 beehives. Aggregated into tropical livestock units (TLU), the average household in the sample owns 1.2 TLUs (median = 1 TLU).

Figure 17. Percent of households owning different livestock species

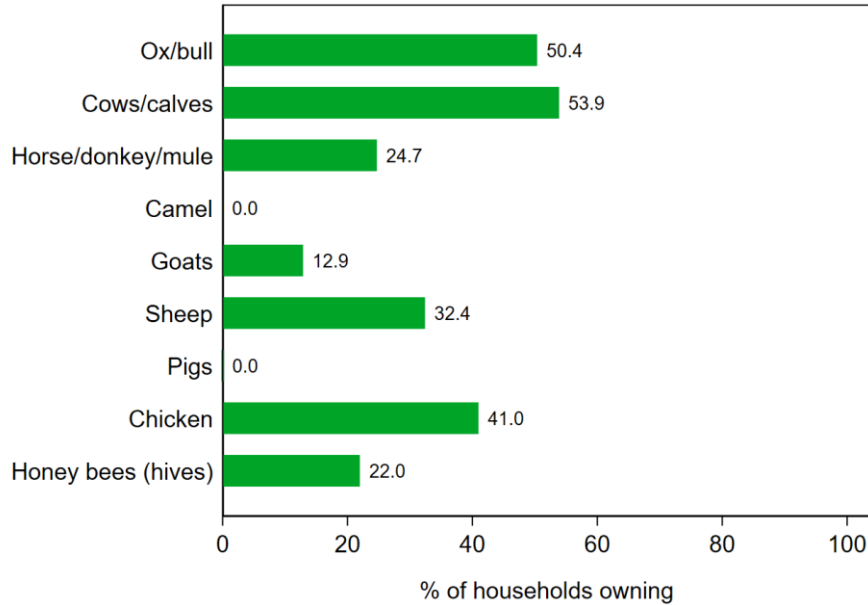
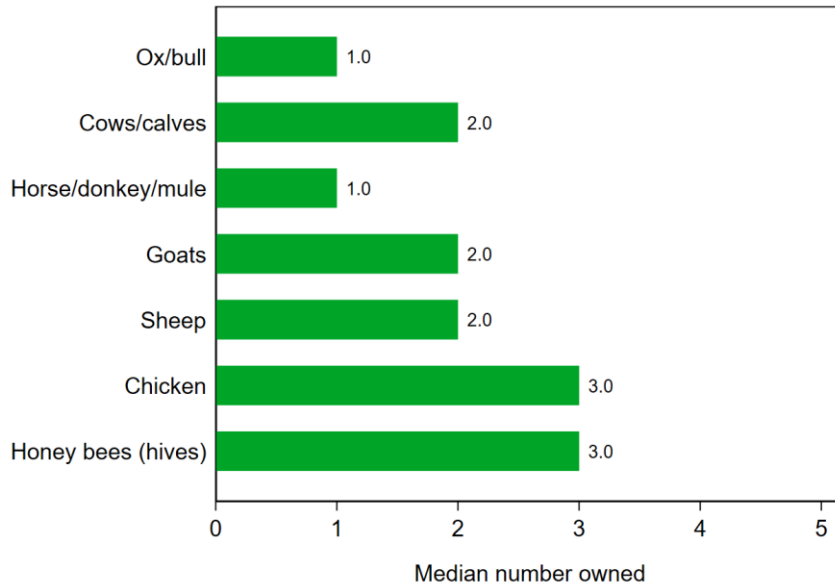


Figure 18. Median number of livestock owned (conditional on owning), by livestock type



Note: Rarely owned livestock types (camel and pigs) omitted from this figure.

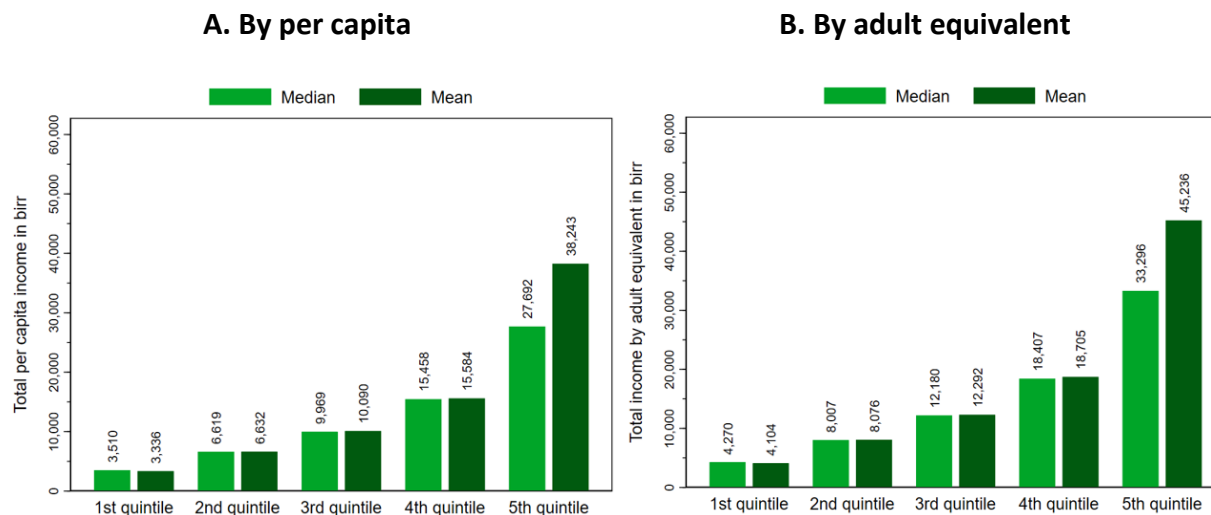
4.5. Incomes

The survey asked households to report incomes generated from crop and livestock production, as well as from enterprises. We also inquired about amounts received through remittances and other income sources, including agricultural and non-agricultural wages, rents, pensions, and social assistance programs. To construct a total per capita income aggregate, we annualized all

incomes and divided them by household size. All incomes are reported in gross terms, except for enterprise incomes, which are reported based on profits. Crop incomes are based on harvested amounts, multiplied by sales prices collected as a part of the survey. As such, they consider harvest that was sold, stored and consumed by the household. Measuring income in this context is notoriously difficult (for more details, see footnote #3). This often results in noisy data that underestimates actual income levels. To mitigate the influence of extreme values, we winsorized all income components at the 95th percentile. Additionally, we report median values alongside means, as medians are less sensitive to outliers.

The mean total per capita annual income in the sample is 14,771 birrs and median 9,968 birrs. The corresponding values based on adult equivalents are 17,674 birrs and 12,170 birrs, respectively. Figure D1 in Appendix D shows limited variation across mean and median values across woredas. We then used these income variables to group households into five quintiles. Figure 20 shows the mean and median incomes for each income quintile. The median annual per capita household income in the bottom quintile (i.e., the poorest 20%) is 3,510 birrs while at the top quintile (the richest 20%) it is 27,693 birrs. The median estimates based on adult equivalents are 4,270 birrs at the bottom quintile and 33,296 birrs at the top quintile.

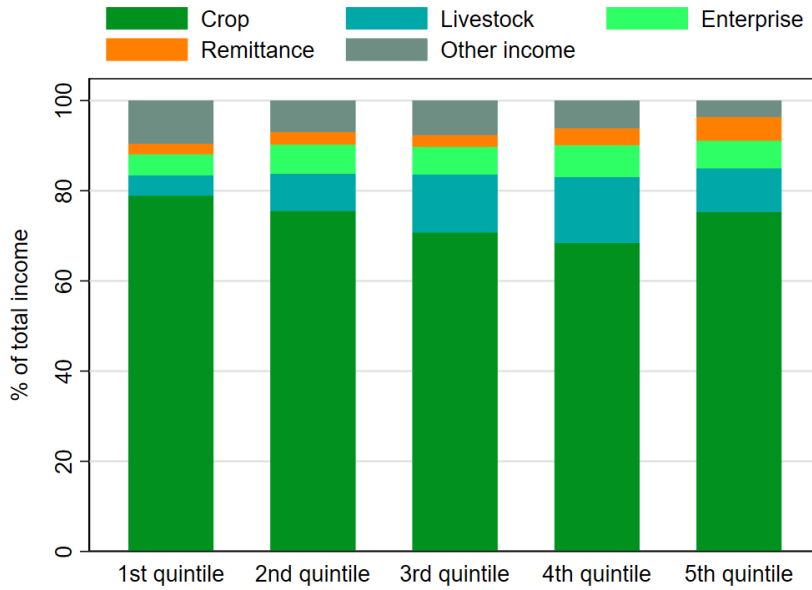
Figure 19. Mean and median total per capita incomes, by income quintile



Note: N = 2,992 households.

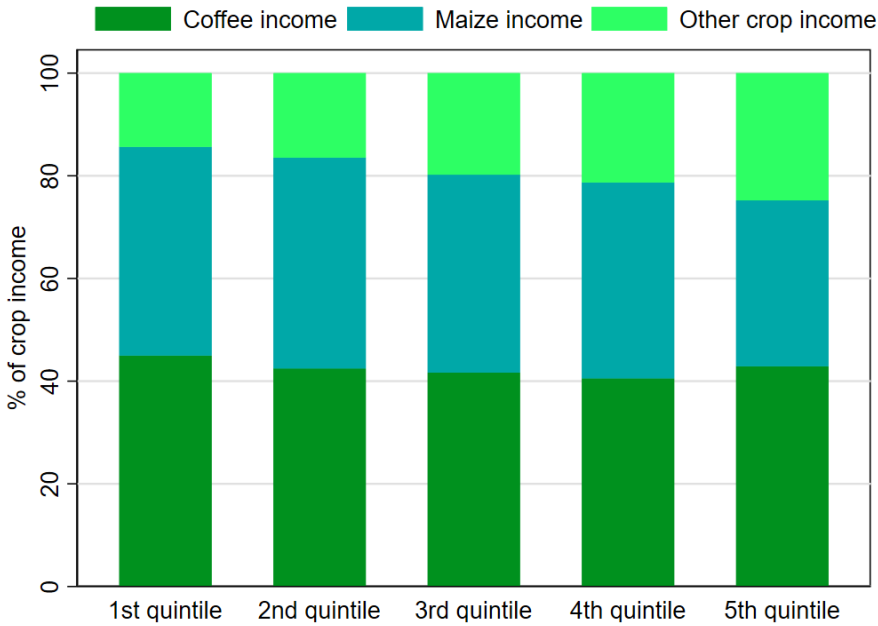
Figure 21 decomposes mean incomes by income component (sources). In each quintile, crop incomes form the lion’s share of the total income, ranging from 78 percent at the bottom quintile to 75 percent at the top quintile. Figure 22 then decomposes crop incomes into income from coffee, maize and other crops. In each quintile, coffee is the most important source of total crop income. However, its relative importance varies across quintiles. At the bottom quintile, 45 percent of the total crop income comes from coffee farming, 41 percent from maize and 14 percent from other crops. At the top quintile, 43 percent of the crop income comes from coffee farming, 32 percent from maize and 25 percent from other crops.

Figure 20. Mean total incomes, by income source and income quintile



N = 2,992 households

Figure 21. Mean crop incomes, by crop and income quintile



N = 2,992 households

Income from livestock sales and products (dairy, eggs, honey) forms about 10 percent of the total income on average. The mean and median livestock income in the sample are 7,997 and 460 birrs, respectively. Livestock incomes are less important at the bottom quintile, forming only

about four percent of the per capita total income. At the top quintile, the corresponding share is 12 percent.

We asked respondents whether anyone in their household was engaged in a non-agricultural income generating activity. About 27 percent of the respondents responded affirmatively with typical non-agricultural activity being petty trading, shopkeeping or a transport business. Less than five percent of these enterprises employed people outside the household. The per capita profits (earnings-minus expenses) estimated by the respondents ranged between 5,000 to 20,000 birr a year (interquartile range) with median profits at 10,000 birr a year. In total, enterprise incomes form only 6.2 percent of the total income (4.4% at the bottom quintile and 7.0% at the top quintile).

While 19 percent of households reported receiving remittances, these were generally small in size, constituting only 3.3 percent of total income on average. About 25 percent of households reported having received wage income in the past 12 months. However, these income sources are relatively unimportant, constituting only 5.8 percent of total income, on average. Meanwhile, few households received income from other sources. Less than three percent received income from renting land or livestock, less than two percent reported pension income, and less than one percent of the households reported receiving income from the Productive Safety Net Program (PSNP) or other social assistance programs.

4.6. Access to loans

We asked respondents whether they had applied for a loan of at least 1,000 birrs in the last 12 months. Two-thirds of the respondents answered "No," citing reasons such as not needing credit or a loan (73% of those who answered "No"), expecting to be rejected (15%), lacking collateral (5%), or thinking that the interest rate was too high (3%) (Table 9). Out of the one-third that responded 'Yes', 90 percent did obtain the loan they had requested. The mean amount borrowed from all sources in the past 12 months was 5,500 birrs. At the time of the interview, the mean outstanding loan amount among these households was 2,800 birrs.

Thirty-five percent of the households that had borrowed reported problems with repayment: they had been unable to make one or more payments in the past 12 months. The typical reason was that the payment was due at a time when the household had less income than usual (Figure 23). Obtaining inputs or goods on credit is rare in this context with only one percent of the households reported obtaining agricultural inputs on credit and seven percent goods on credit (Table 9).

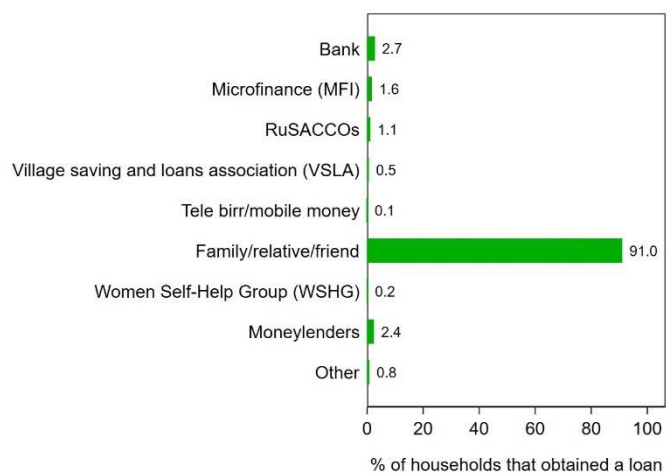
Table 9. Monetary loan, inputs credit and goods credit in the last 12 months

	N	%/mean
Monetary credit (loan)		
% applied for loan in the last 1 year	2,995	34%
% obtained the requested loan	1,011	90%
Total loan from all sources (in 1 year)	913	5,509
Total outstanding loan	913	2,793
% unable to make one or more payments	913	35%
Agricultural input credit		
% obtained inputs on credit in the last 1 year	2,995	1%
Goods on credit		
% obtained goods on credit in the last 1 year	2,995	7%

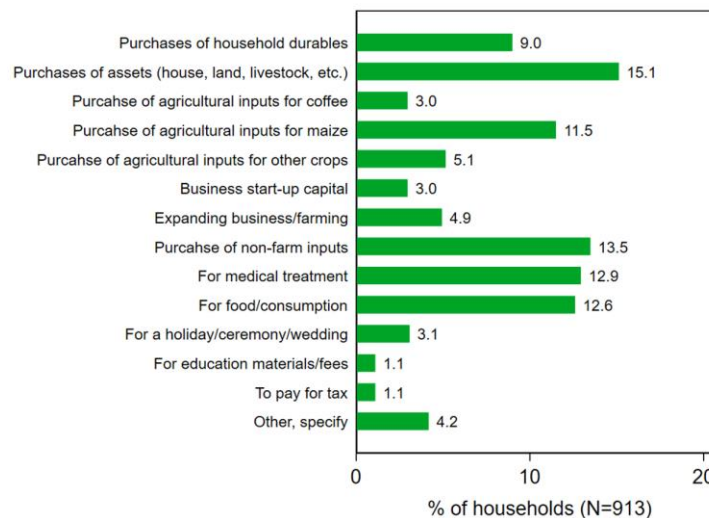
As shown in Figure 23, nine out of ten of the loans were obtained from friends and relatives. Only about five percent of the respondents reported obtaining their loan from a bank or micro-finance institution. Loans were used for a wide variety of purposes (Figure 23): purchasing large assets, such as houses, land or livestock, farm or non-farm inputs, for medical treatment, to meet consumption needs, and so on.

Finally, we asked respondents whether they would be able to secure a 10,000-birr loan if they needed it. More than 60 percent of the households responded negatively and 28 percent affirmatively. The most likely source for this loan would be friends and relatives (Figure 23).

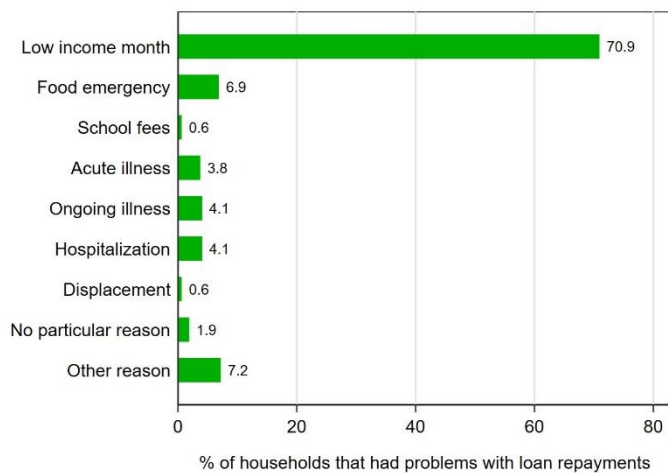
Figure 22: Source of loan, purpose, reasons for delayed payment, and potential lenders



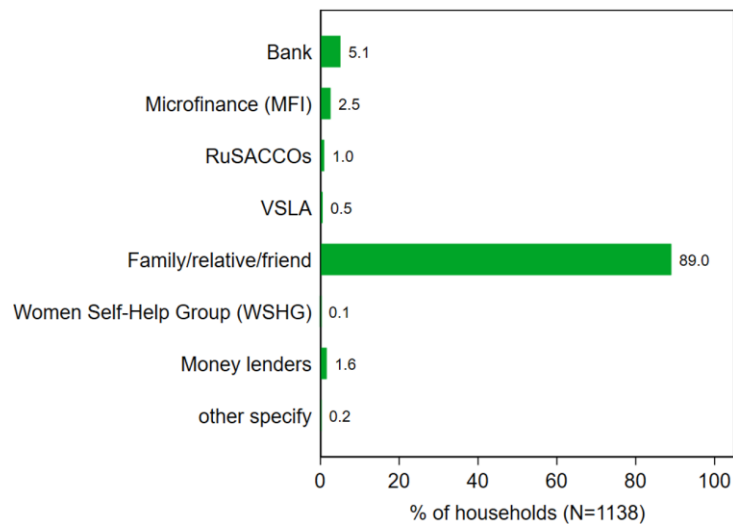
(a) Loan source



(b) Loan purpose



(c) Reason for being unable to repay



(d) Potential lenders

4.7. Consumption

Following recommended practices (Deaton & Grosh, 2000), the baseline survey questionnaire asked households to report on their consumption expenditures over a specific time interval. The consumption-expenditure module had three sub-modules: 1. Household food consumption over the past 7 days; 2. Household expenditure on consumables over the past 1 month; and 3. Household expenditures on durables and services over the past 12 months.

We used these data to construct food, non-food and total consumption aggregates (Deaton and Zaidi 2002). To this end, we converted household food consumption measured in quantity terms to birr terms using the price data collected as a part of the survey.⁸ Once all consumption components were measured in birr terms, the consumption aggregate was formed of the three consumption components listed above. All consumption amounts were transformed into daily terms by dividing the reported amount by the number of days in the recall period. Daily household consumption expenditures were reported on a per capita basis by dividing the consumption expenditure by household size.

To estimate the prevalence and depth of poverty based on the international poverty line, we converted the per capita consumption expenditure values measured in birr terms to 2017 Purchasing Power Parity (PPP) dollars. To do so, we used food consumer price index (CPI) series provided by the Ethiopian Statistical Service to deflate the consumption estimates to 2017 prices.⁹ We then converted the consumption-expenditure values expressed in 2017 prices to PPP terms using the PPP conversion rate for individual consumption expenditure by households, obtained from the International Comparison Program (ICP). The 2017 \$PPP exchange rate for Ethiopia is 8.21.

Household consumption and poverty estimates are summarized in Table 10. The mean daily household per capita food consumption expenditures in this sample is 72.42 birr and the mean daily per capita non-food expenditures is 11.26 birr. Therefore, the average household in our sample spends 86 percent of its total budget on food, indicative of the high level of deprivation in this sample of households. Adding the mean food and non-food consumption expenditures amounts to a mean total per capita consumption expenditure of 83.68 birr per day. This then translates into 2.33 \$PPP, which is only marginally higher than the international extreme poverty line: \$2.15 \$PPP.

⁸ The study used a price opinion survey (asking households to estimate prices) instead of the usual method of sending enumerators to local food markets. This is because markets often lack certain food items, especially meat and milk in rural areas. While this survey method is relatively uncommon, previous experimental work suggests it yields accurate price data (Gibson and Rozelle 2003). Households were asked to estimate prices for a randomly chosen subset of 10 food items out of a list of 65 food items included in the food consumption module. Price estimates were converted to a standard unit and then averaged at the woreda level using the median price to reduce outlier influence.

⁹ For the average household in our sample food expenditures form 86 percent of the total consumption basket. Therefore, we used food consumer price index (CPI) estimates to deflate the estimates to 2017 prices.

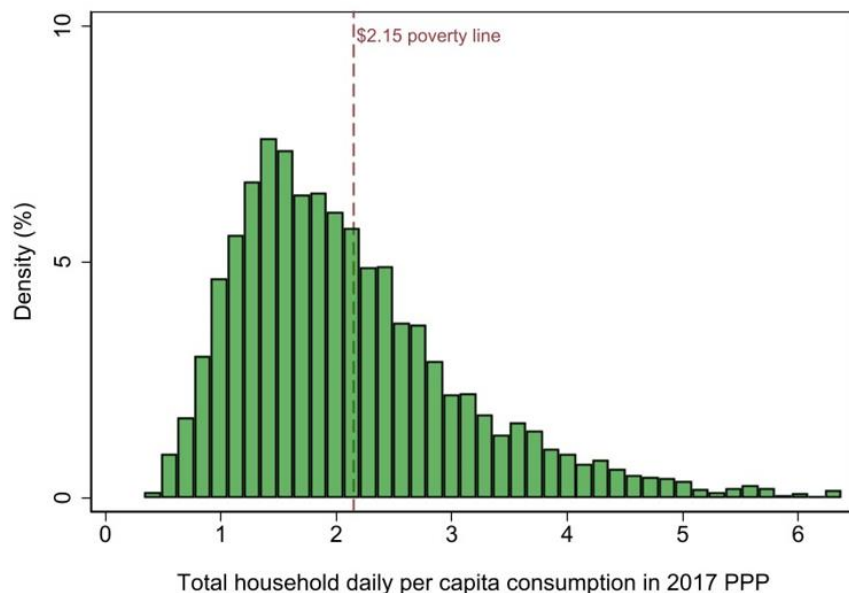
More than half (60%) of the individuals originate from households for which the daily per capita consumption is below 2.15 \$PPP. The poverty gap index tells us that the consumption level of the average poor in the sample is 31 percent of the poverty line, implying that a 0.67 \$PPP (= 0.31 * 2.15 \$PPP) increase in daily per capita consumption would be required to lift the average poor person in the sample above the poverty line. Figure 24 overlays the distribution of household daily per capita consumption on the 2.15. We see that many individuals are located just below or just above the poverty line, indicating that even relatively small changes in consumption levels may induce large shifts in poverty headcount estimates. Finally, the risk of falling below the poverty line increases with household size (Figure 25). The opposite is true for land holdings size, though the slope is not as steep (Figure 26).

Table 10. Household consumption and poverty indicators

Outcome	N	mean
Daily per capita food consumption-expenditures in birr	2,992	72.42
Daily per adult equivalent food consumption-expenditures in birr	2,992	87.60
Daily per capita non-food consumption-expenditures in birr	2,992	11.26
Daily per adult equivalent non-food consumption-expenditures in birr	2,992	13.54
Daily per capita total consumption-expenditures in birr	2,992	83.68
Daily per adult equivalent total consumption-expenditures in birr	2,992	101.13
Daily per capita total consumption-expenditures in 2017 \$PPP	2,992	2.33
Daily per adult equivalent total consumption-expenditures in 2017 \$PPP	2,992	2.82
Poverty headcount, based on 2.15 \$PPP poverty line (*)	16,364	60.28
Poverty gap index, based on 2.15 \$PPP poverty line (*)	9,865	31.39

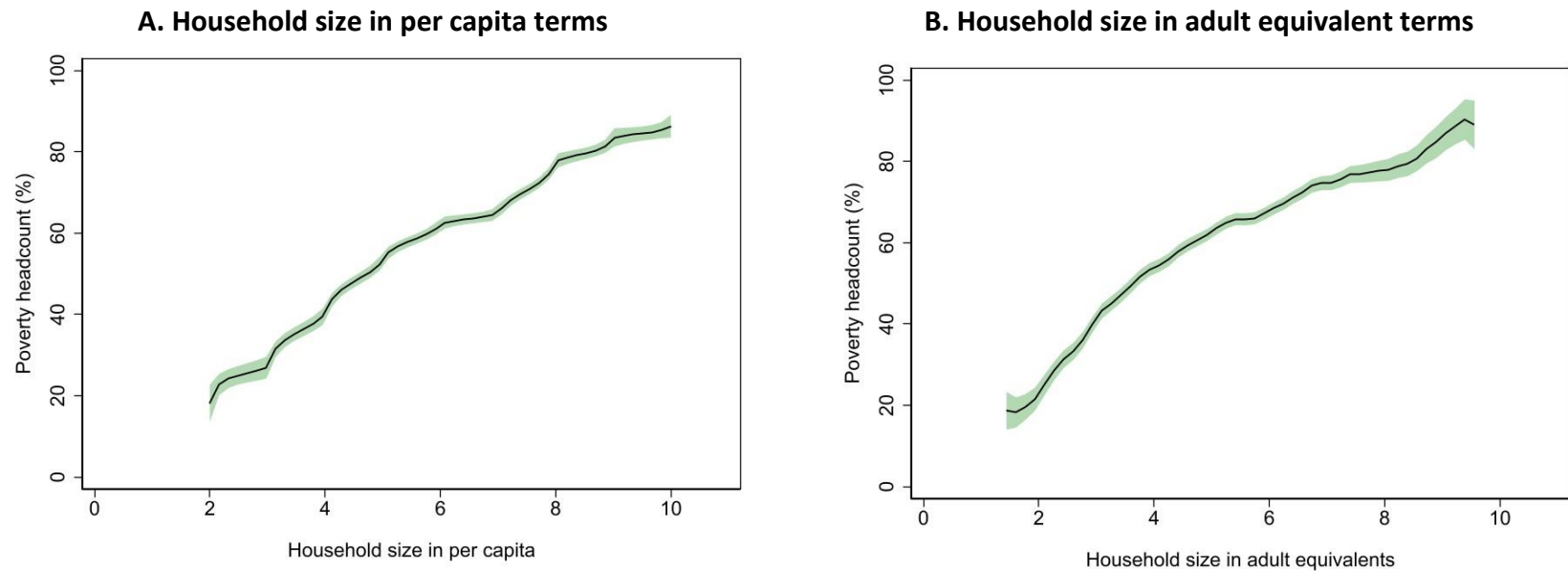
Note: (*) Individual level measure; means and CIs estimated using household size as the frequency weight. Household roster data were incomplete for three households and therefore the number of household level observations is 2,992 instead of 2,995.

Figure 23. Distribution of household daily per capita consumption in 2017 PPP



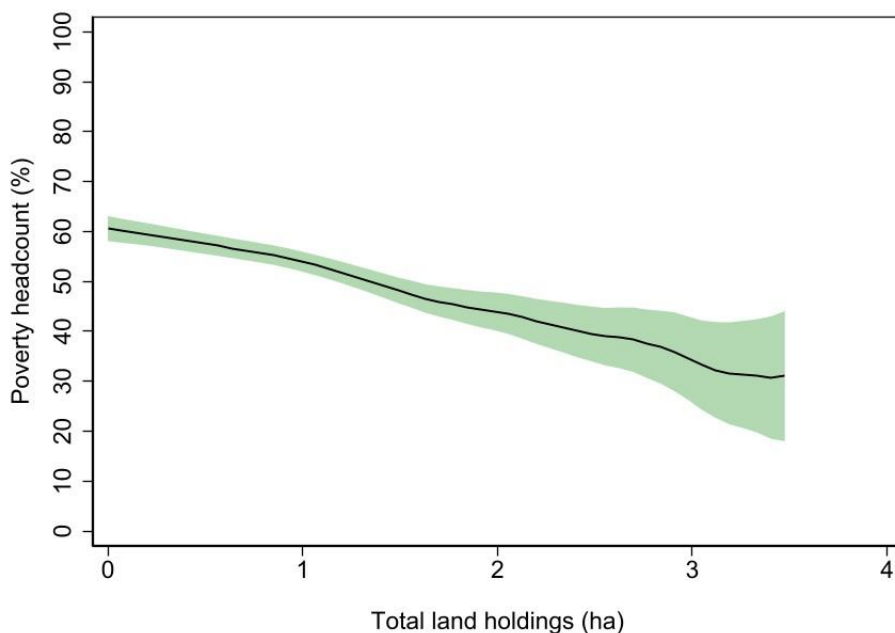
Note: Household consumption estimates are weighted by household size. The horizontal axis is truncated at the 99th percentile of the consumption distribution.

Figure 24. Risk of extreme poverty, by household size



Note: Local polynomial regression. Shaded lines represent 95 percent-confidence intervals. Households of very small size (less than 1) and very large size (greater than 10) have been excluded from the sample because they could disproportionately influence the regression at the distribution tails due to fewer observations.

Figure 25. Risk of extreme poverty, by land size



Note: Local polynomial regression. Shaded lines represent 95 percent-confidence intervals. The horizontal axis is truncated at the 99th percentile of the land holdings distribution.

4.8. Food security

We use two indicators to assess household food security: the Food Insecurity Experience Scale (FIES) developed by Ballard, Kepple, and Cafiero (2013) and household diet diversity score (HDDS) developed by Swindale and Bilinsky (2006).

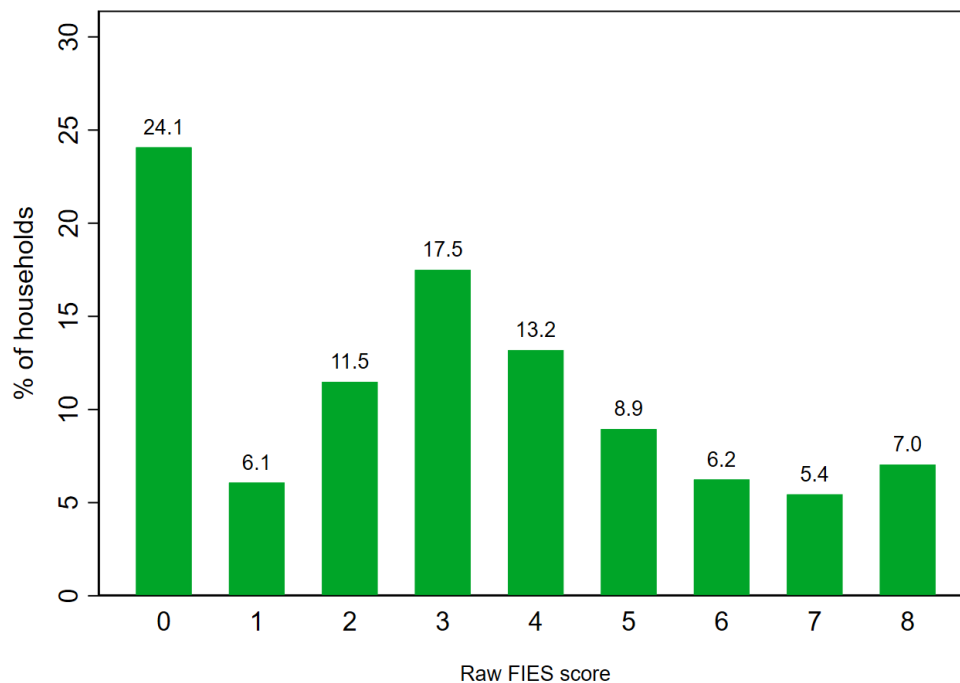
The FIES module asks about a household's access to food within a specific recall period, which in this survey was set at 12 months. Table 11 details the questions asked and the percentage of households answering affirmatively to each one. The severity of food insecurity increases as one moves down the list of questions, which explains why the percentage of households responding positively to the question decreases.

We assigned a value 1 to each positive response and added them up to construct a (raw) FIES score. The distribution of these scores is presented in Figure 27. Following the guidelines established by the Food and Agricultural Organization of the United Nations, a score of zero (indicating "No" responses to all questions) signifies food security at the household level. Scores ranging from 1 to 3 suggest mild food insecurity, scores between 4 and 6 indicate moderate food insecurity, and scores exceeding 6 are indicative of severe food insecurity. Based on this indicator, 24 percent of the households were food secure, 35 percent mildly food insecure, 28 percent moderately food insecure and 12.5 percent severely food insecure.

Table 11. Food Insecurity Experience Scale questions

Category	%
FIES categories:	
Food secure	24.1
Mildly food insecure	35.1
Moderately food insecure	28.4
Severely food insecure	12.5
Share of households responding positively to FIES questions:	
During the past 12 months...	
Was there a time when you or others in your household were worried you would not have enough food to eat because of a lack of money or other resources?	62.8
Was there a time when you or others in your household were unable to eat healthy and nutritious food because of a lack of money or other resources?	69.2
Was there a time when you or others in your household ate only a few kinds of foods because of a lack of money or other resources?	65.8
Was there a time when you or others in your household had to skip a meal because there was not enough money or other resources to get food?	30.0
Was there a time when you or others in your household ate less than you thought you should because of a lack of money or other resources?	41.1
Was there a time when your household did not have food because of a lack of money or other resources?	18.7
Was there a time when you or others in your household were hungry but did not eat because there was not enough money or other resources for food?	15.4
Was there a time when you or others in your household went without eating for a whole day because of a lack of money or other resources?	7.9

Figure 26. Distribution of the Food Insecurity Experience Scale (FIES) scores



Next, we look at HDDS, which groups food consumption into 12 food groups, listed in Table 12. Consumption from each food group gives a point, yielding a count HDDS variable ranging between 1 and 12. The mean HDDS in this sample is 7.6 and the median is 8. Virtually all households consumed Cereals, Vegetables, Pulses, Oils and Miscellaneous foods while the consumption of animal source foods was relatively less common.

Table 12. Household Diet Diversity Score (HDDS)

Category	Value
Mean HDDS	7.58
Median HDDS	8.00
Percent consuming from:	
Cereals	99.83
Root and tubers	73.19
Vegetables	99.97
Fruits	66.91
Meat, poultry, offal	16.99
Eggs	20.30
Fish and seafood	0.23
Pulses/legumes/nuts	99.70
Milk and milk products	29.12
Oil/fats	99.47
Sugar/honey	52.35
Miscellaneous	100.00

Note: N = 2,995 households.

4.9. Women's self-help groups

About 9 percent (271/2873) of the women in the sample reported belonging to a women's self-help group (WSHG). A woreda-level disaggregation reveals that most of these groups operated in the Shabe Sombo woreda, where one-third of the sampled women belonged to a WSHG (Figure 28). The corresponding percentages in the other three woredas ranged between two and six percent.

A typical WSHG has about 30 women, and about two-thirds of the groups organize training for their members (Table 13). Out of these groups, 44 percent had organized training on agriculture, 56 percent on business management, 71 percent on financial management, and 70 percent on life skills. Other WSHG activities include the provision of savings services (70 percent of the WSHGs) and loan/credit services (39 percent). Other income-generating activities (3 percent of the WSHGs) or revolving loans/funds (less than 2 percent) were rare.

Figure 27. Percent of women belonging to a women’s self-help group, by woreda

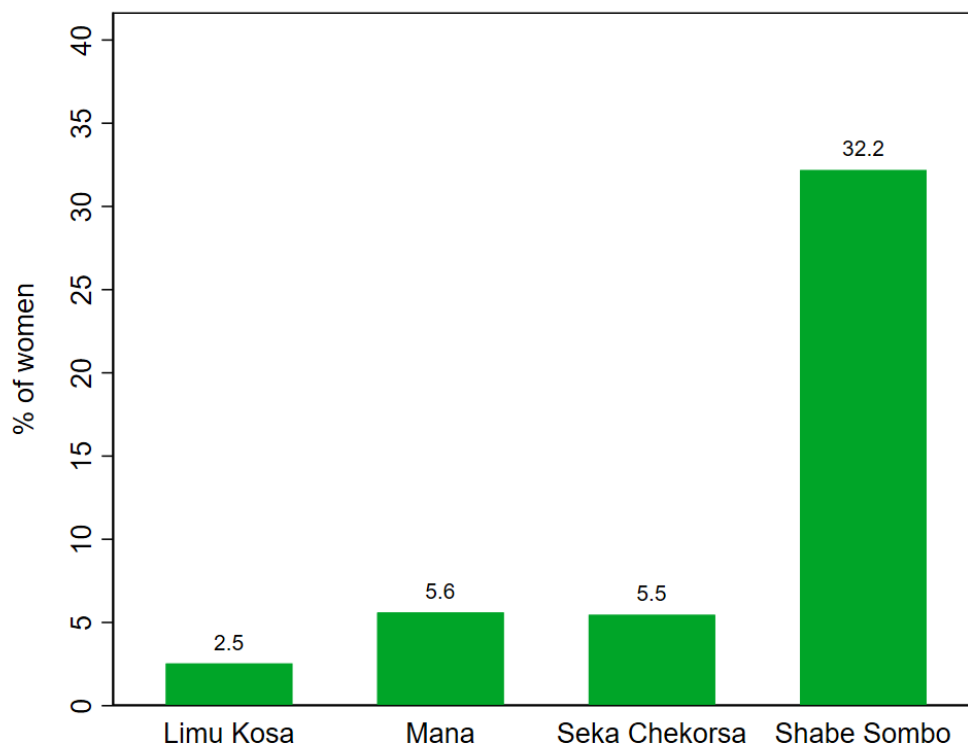


Table 13. Women’s self-help group activities

Activity	%
Savings services	70.9
Loan/credit services	39.5
Any training	63.1
Theme of the training if organized:	
Agriculture	44.4
Business management	56.1
Financial management	70.8
Life skills training	70.2
Other income-generating activities	3.3
Revolving loans/funds	1.9

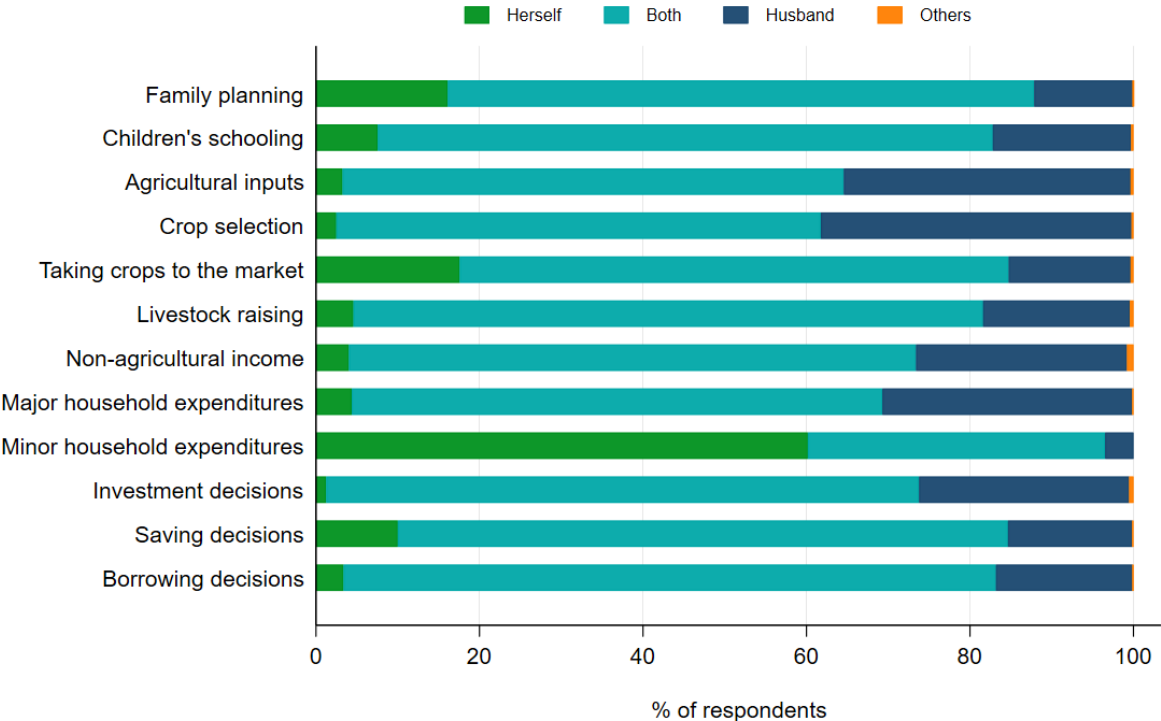
N = 271 women who belong to women’s self-help groups

Finally, the survey instrument included questions about the role of women in decision-making across various domains, such as family planning, children's schooling, agricultural and non-

agricultural production, and others.¹⁰ The complete list of the 12 different domains is detailed in Figure 29. In nearly all domains, the most frequent response was that decisions are made jointly by the husband and wife. For instance, 72 percent of respondents indicated that decisions regarding family planning are made jointly, 16 percent of women stated that they make these decisions independently, and in 12 percent of cases, the husband is the sole decision-maker. While decisions concerning agricultural inputs and crops are typically made together, in more than 35 percent of households, the husband decides. The only domain where women predominantly make decisions alone is concerning minor household expenditures; in 60 percent of households, the female spouse is the decision-maker.

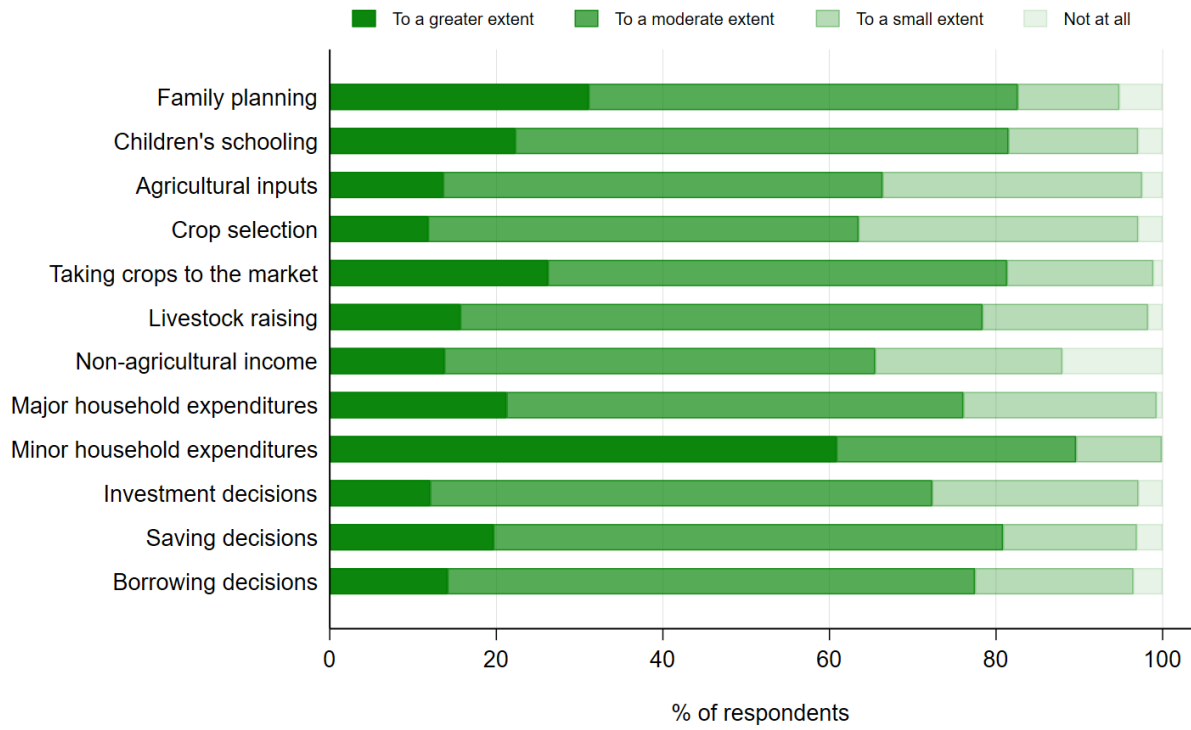
Across all domains, 60 percent of the women feel empowered to make their own personal decisions (Figure 30). Domains where the perception of decision-making power was weakest included those related to agricultural inputs, crop selection, and non-agricultural income. In these areas, more than 30 percent of the women indicated that they could influence decisions not at all or to only a small extent.

Figure 28. Decision making within households across different domains



¹⁰ In 88 percent of the interviews, the female spouse was alone with the enumerator when responding to these questions. In about nine percent of the time, children or other adult females were present. In three percent of the time, an adult male was present when this module was administered. Female-headed households have been omitted from this analysis.

Figure 29. Degree to which women feel they can make their own personal decision



5. Summary and conclusions

The objective of this report is to present the baseline findings from the impact evaluation of the Improving Incomes of Coffee Farmers (IICoF) project in the Jimma zone of Ethiopia. The project aims to improve the income and livelihoods of smallholder farmers who rely on coffee as their primary source of income, while also recognizing the importance of maize in ensuring food security. The bundled intervention attempts to address multiple technology adoption constraints in coffee and maize farming with the goal of improving farm productivity and household incomes. The evaluation aims to generate valuable insights and lessons that can be applied to scaling up similar efforts in Ethiopia and other low-income country settings.

The evaluation employs a randomized controlled trial (RCT) to assess the causal effect of the interventions on the adoption of good agricultural practices, productivity, and income. The evaluation will also assess potential synergies between the interventions, provide insights into the pathways that help explain the realized impacts, and potential heterogeneity of impacts along economic and social dimensions.

The baseline survey was conducted in December 2023 and January 2024 in the four study districts (woredas). A sample of 2,995 households grouped into 750 clusters (development groups) was included in the survey. The inclusion criteria required that households were farming coffee at the time of the baseline survey.

Analysis of the baseline data shows that the sample for this evaluation is characterized by a high level of food insecurity and poverty. Poverty headcount based on the \$PPP 2.15 international poverty line is 60 percent, and more than 40 percent of the households are either moderately or severely food insecure based on FAO's Food Insecurity Experience Scale (FIES).

Household size is a strong predictor of deprivation risk in this context. The average household in the sample is formed of five members (4.5 in terms of adult equivalents, AEU). However, this average masks considerable heterogeneity: the smallest households have just one member (0.74 AEU) whereas the largest one has 18 members (14.9 AEU). Less than nine percent of the households are female headed. The mean age of the household head is 45 years and 43 percent of the household heads are illiterate (cannot read or write). Access to main grid electricity is rare and most households (79%) use solar energy to light their houses after sunset.

The mean (median) agricultural land area operated by sample households is approximately 0.97 ha (0.75 ha), out of which 0.51 ha (0.31 ha) is allocated to coffee farming, 0.27 ha (0.25 ha) to maize farming and 0.19 ha (0.06 ha) to other crops. The average farmer in our sample produced 157 kg (median: 86 kg) of green beans during the 2023 season and the mean yield was 348 kg/ha (median: 262 kg/ha). The average maize farmer reported producing 694 kg of maize (median: 500 kg), and the mean yield was about 2,000 kg per hectare (median: 1,714 kg/ha). While a substantial share of coffee produced is sold, most of the maize is used for home consumption.

The survey questionnaire included a quiz testing farmers' knowledge of recommended coffee farming practices. This quiz revealed that farmers have a good understanding of recommended practices for shade management and stumping methods. However, there were considerable

knowledge gaps when it came to preparation of holes for planting new coffee trees. We also asked about the adoption of these practices. Over 80 percent of households reported using shade to protect their coffee trees from excessive heat, 66 percent practiced integrated disease management and nearly 60 percent had weeded their coffee farms in the past year. However, only about 20 percent followed recommended soil fertility management practices. Just 10 percent had rejuvenated their coffee trees (stumping) in the last three years, and even fewer kept records, possibly due to low literacy levels.

Regarding maize farming, the farmers demonstrated a solid grasp of recommended practices for ploughing, NPS fertilizer application, and FAW control methods. However, their understanding of the recommended practices for UREA application, inter-cultivation and row planting was generally limited. In terms of actual adoption, 80-90 percent of maize farmers reported ploughing more than three times before sowing, and applied FAW control method when there was incidence or frequently inspected their maize field. Nonetheless, most did not apply the recommended per-hectare amounts of fertilizers.

Overall, the baseline survey describes a context that supports the project's underlying assumptions. It highlights an urgent need to improve smallholder incomes and livelihoods to reduce poverty and food insecurity. The project's theory of change posits that these objectives can be achieved by promoting improved agronomic practices, combined with better access to credit and production inputs. The follow-up surveys permit us to rigorously evaluate the effectiveness of this approach.

References

- Abate, Gashaw T, Tanguy Bernard, Simrin Makhija, and David J Spielman. 2023. "Accelerating technical change through ICT: Evidence from a video-mediated extension experiment in Ethiopia." *World Development* 161:106089.
- Abate, Gashaw T, Kalle Hirvonen, and Belay Mohammed. 2021. HereWeGrow's Jimma Coffee Program (JCP): Qualitative Assessment Report at Baseline. Unpublished mimeo.
- Abate, Gashaw T, Kalle Hirvonen, and Belay Mohammed. 2022. HereWeGrow's Jimma Coffee Program (JCP). Baseline Report (2022 Cohort). Unpublished mimeo.
- Ballard, Terri J, Anne W Kepple, and Carlo Cafiero. 2013. The food insecurity experience scale: development of a global standard for monitoring hunger worldwide. *FAO Technical Paper Version 1.1*.
- Berhane, Guush, Gashaw Tadesse Abate, and Abdulazize Wolle. 2022. Agricultural intensification in Ethiopia: Patterns, trends, and welfare impacts. *IFPRI Discussion Paper* 2150.
- Carletto, Gero, Marco Tiberti, and Alberto Zezza. 2022. "Measure for measure: Comparing survey based estimates of income and consumption for rural households." *The World Bank Research Observer* 37 (1):1-38.
- Carril, Alvaro. 2017. "Dealing with misfits in random treatment assignment." *The Stata Journal* 17 (3):652-667.
- Deaton, Angus, and Margaret Grosh. 2000. "Consumption." In *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of Living Standards Measurement Study*, edited by Margaret Grosh and Paul Glewwe, 91-133. Washington D.C.: World Bank.
- Deaton, Angus, and Salman Zaidi. 2002. *Guidelines for constructing consumption aggregates for welfare analysis*. Vol. 135: World Bank Publications.
- Deininger, Klaus, Calogero Carletto, Sara Savastano, and James Muwonge. 2012. "Can diaries help in improving agricultural production statistics? Evidence from Uganda." *Journal of Development Economics* 98 (1):42-50.
- Duflo, Esther, Daniel Keniston, Tavneet Suri, and Céline Zipfel. 2023. Chat over coffee? diffusion of agronomic practices and market spillovers in Rwanda. National Bureau of Economic Research.
- GoE. 2020. A Homegrown Economic Reform Agenda: A Pathway to Prosperity. Addis Ababa: Government of Ethiopia (GoE).
- Hoppenbrouwers, M. and R. van Haaften. 2023. HereWeGrow 2019 Coffee Yield Survey Ethiopia: Report. Addis Ababa, Ethiopia.
- McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99 (2):210-221.
- Sahn, David E, and David Stifel. 2003. "Exploring alternative measures of welfare in the absence of expenditure data." *Review of income and wealth* 49 (4):463-489.
- Shikuku, Kelvin Mashisia. 2019. "Information exchange links, knowledge exposure, and adoption of agricultural technologies in northern Uganda." *World Development* 115:94-106.
- World Bank. 2021. "Ethiopian Farmers Triple Coffee Yields with Sustainable Tree Stumping." 20 April 2021. <https://www.worldbank.org/en/news/feature/2021/04/20/ethiopian-farmers-triple-coffee-yields-with-sustainable-tree-stumping>.

Appendix

Appendix A: Randomization and sampling protocol

Introduction

The unit of randomization and treatment is the development group (DG). The randomization of DGs was conducted in two stages. First, the project area covers 75 kebeles out of which 55 kebeles are randomly selected to be part of the project and the remaining 20 kebeles serve as pure control kebeles. Second, we randomized DGs into study arms. Within the 20 pure control kebeles, 125 DGs were randomly selected to be part of the study. Within the 55 treatment kebeles, 625 DGs were randomly selected and allocated to 5 treatment arms. After the DG selection, we selected 4 households from each DG to be part of the household sample used in the baseline, midline and endline surveys.

Household listing

A full household listing was conducted in all 75 project kebeles in October 2023. The listing collected information about the woreda, kebele and village in which the household is located, as well as the name of the DG it belongs to. In addition, the listing form included basic information about the household: name, sex and contact details of the household head and spouse. We also asked households if they cultivated coffee. The first listing covered 53,047 households in 1,949 DGs. An additional listing of 1,949 households in 35 DGs was conducted later.

Eligibility criteria

Our sampling frame is formed of:

- Households that reported cultivating coffee during the listing.
- Households belonging to DGs in which the listing was conducted during the first pass (i.e., we exclude the 35 DGs in which the listing was conducted later).
- Households belonging to DGs that have at least 7 other coffee cultivating households.

These eligibility criteria only pertain to the sampling frame. Households that are excluded from the sampling frame are eligible to receive treatment. To ensure this, we allocated all households, irrespective of whether they were selected to the sample, into the five arms.

1st stage randomization: selecting kebeles into treatment and (pure) control

The project area is formed of 75 kebeles in 4 woredas. In the first stage randomization, we randomly selected 55 kebeles to be part of the treatment area. To prevent an unequal distribution of kebeles across woredas, we used a block randomization approach so that between 72 and 75 percent of all kebeles in a woreda were allocated to treatment. Table A1 lists the kebeles selected for treatment and (pure) control.

Table A1. Number of kebeles, by treatment status and woreda

Woreda	Pure control kebeles	Treatment kebeles	% selected for treatment
Limu Kosa	7	18	72
Mana	5	14	74
Seka Chekorsa	5	14	74
Shabe Sombo	3	9	75
Total	20	55	73

2nd stage randomization: allocating DGs into treatment arms

Pure control kebeles: Within the 20-pure control kebeles, we allocated 125 DGs to be part of the sample. We used block randomization methods to ensure balance across kebeles with misfits allocated globally, Carril (2017). Table A2 shows the number of sampled DGs per kebele in Pure control areas. In each kebele, between 19 and 26 percent of the DGs were randomly selected to be included in the study sample.

Table A2. Pure control areas: number of DGs selected, by kebele

	Selected	Not selected	% selected
Limu Kosa woreda:			
Arangema	5	14	26
Chafe Ifeta	5	20	20
Dangaja Sole	6	21	22
Gena Dujuma	2	8	20
Harewa Jimate	5	17	23
Tencho	6	21	22
Walensu	6	21	22
Mana woreda:			
Gube Bosoka	6	21	22
Hunda Toli	6	17	26
Lemi Lelisa	11	38	22
Sombo Mana	8	30	21
Urgeyi	9	30	23
Seka Chekorsa woreda:			
Atro Sufa	4	17	19
Budo Kera	4	14	22
Debo Yaya	7	25	22
Ilke Tunjo	5	21	19
Lilu Ceha	8	25	24
Shabe Sombo woreda:			
Hane Da'O	8	33	20
Leku Migira	6	24	20
Shebedeso	8	26	24
Total	125	443	

Treatment area kebeles: Within the 55-treatment area kebeles, we first selected 625 eligible DGs to be part of the sample while the remaining 742 eligible DGs were marked as non-sample DGs. Using block randomization methods, these 625 DGs were balanced across kebeles. Table A3 summarizes the selection across the treatment kebeles. In each treatment kebele, between 38 and 56 percent of the DGs were selected to be included in the study sample.

Table A3. Treatment areas: number of eligible DGs selected, by kebele

	Selected	Not selected	% selected
Limu Kosa woreda:	201	237	46
Adis Limat	5	8	38
Chancho Gashe	10	14	42
Dembi Gebena	10	13	43
Dire Togo	11	15	42
Gena Dambi	11	12	48
Goboze Wabe	11	12	48
Harawa Gatira	25	28	47
Harawa Kata	12	13	48
Kechewo Tirtira	9	12	43
Kela Gabisa	13	17	43
Kossa Geshe	11	11	50
Leku Chime	10	11	48
Liban Gonde	11	11	50
Mendera	11	12	48
Mito	8	11	42
Suntu	9	11	45
Tanebo Lalo	12	13	48
Walake Sombo	12	13	48
Mana woreda:	172	209	45
Bilida	14	16	47
Dawa	11	14	44
Doyo Bikila	12	13	48
Doyo Toli	16	17	48
Gube Muleta	10	13	43
Haro	12	13	48
Hunda Keneni	17	18	49
Itisa Guda	11	14	44
Kela Guda	12	15	44
Kentiri	17	22	44
Kore Lalisa	9	12	43
Meti	8	11	42
Saye Bontu	10	14	42
Somodo	13	17	43
Seka Chekorsa woreda:	140	167	46

	Selected	Not selected	% selected
Andode Alaga	11	15	42
Buyo Kechemma	13	15	46
Gepa Sadan	5	7	42
Gibe Boso	9	10	47
Gudo Daka	16	20	44
Kora Waco	3	5	38
Kusaro	5	4	56
Nase	9	12	43
Sekela Genefo	18	23	44
Sentema Goro	11	11	50
Shashamene	13	14	48
Siba Kake	7	7	50
Ushane Buyo	6	6	50
Wokitu Medalu	14	18	44
Shabe Sombo woreda:	112	129	46
Angecha	12	15	44
Anja Gombo	14	15	48
Gembo Migira	7	10	41
Halo Godante	19	21	48
Halo Soboka	18	20	47
Sabaqo Dabiye	11	11	50
Sebaka Wala	9	9	50
Wala Kela	10	13	43
Yanga Dogoma	12	15	44
Total	625	742	46

After this, the 625 DGs selected to be part of the sample used in this study were allocated to 5 treatment arms. We used block randomization methods (with misfits allocated globally) to ensure equal representation of treatment arms within each kebele. Table A4 shows the number of DGs by treatment arm and by kebele.

Table A4. Treatment areas (study DGs): number of DGs, by treatment arm and kebele

Treatment arm:	1	2	3	4	5	Total
Limu Kosa woreda:						
Adis Limat	1	1	1	1	1	5
Chancho Gashe	2	2	2	2	2	10
Dembi Gebena	2	2	2	2	2	10
Dire Togo	2	2	2	3	2	11
Gena Dambi	2	2	2	2	3	11
Goboze Wabe	2	3	2	2	2	11
Harawa Gatira	5	5	5	5	5	25
Harawa Kata	3	2	3	2	2	12

Treatment arm:	1	2	3	4	5	Total
Kechewo Tirtira	2	2	1	2	2	9
Kela Gabisa	2	2	3	3	3	13
Kossa Geshe	2	3	2	2	2	11
Leku Chime	2	2	2	2	2	10
Liban Gonde	3	2	2	2	2	11
Mendera	2	2	3	2	2	11
Mito	1	2	1	2	2	8
Suntu	2	1	2	2	2	9
Tanebo Lalo	3	3	2	2	2	12
Walake Sombo	2	2	3	3	2	12
Mana worda:						
Bilida	3	3	3	2	3	14
Dawa	2	2	2	3	2	11
Doyo Bikila	2	3	2	2	3	12
Doyo Toli	4	3	3	3	3	16
Gube Muleta	2	2	2	2	2	10
Haro	2	2	3	3	2	12
Hunda Keneni	3	4	3	3	4	17
Itisa Guda	3	2	2	2	2	11
Kela Guda	2	2	3	3	2	12
Kentiri	3	4	3	3	4	17
Kore Lalisa	2	1	2	2	2	9
Meti	2	2	2	1	1	8
Saye Bontu	2	2	2	2	2	10
Somodo	2	3	2	3	3	13
Seka Chekorsa worda:						
Andode Alaga	3	2	2	2	2	11
Buyo Kechema	2	2	3	3	3	13
Gepa Sadan	1	1	1	1	1	5
Gibe Boso	2	2	2	2	1	9
Gudo Daka	3	3	3	3	4	16
Kora Waco	1	1	1	0	0	3
Kusaro	1	1	1	1	1	5
Nase	2	2	1	2	2	9
Sekela Genefo	3	3	4	4	4	18
Sentema Goro	2	3	2	2	2	11
Shashamene	3	2	3	3	2	13
Siba Kake	1	2	1	1	2	7
Ushane Buyo	2	1	1	1	1	6
Wokitu Medalu	2	3	3	3	3	14
Shabe Sombo worda:						
Angecha	3	2	3	2	2	12
Anja Gombo	3	3	2	3	3	14

Treatment arm:	1	2	3	4	5	Total
Gembo Migira	1	1	2	2	1	7
Halo Godante	4	4	4	3	4	19
Halo Soboka	3	4	3	4	4	18
Sabaqo Dabiye	3	2	2	2	2	11
Sebaka Wala	1	2	2	2	2	9
Wala Kela	2	2	2	2	2	10
Yanga Dogoma	3	2	3	2	2	12
Total	125	125	125	125	125	625

We followed the same approach to allocate the 742 eligible DGs as well as 145 non-eligible¹¹ DGs not selected to be included in the study sample into the 5 treatment arms. Therefore, the total number of non-sampled DGs in the treatment area is 887.

Selecting households for the surveys

The total sample used in this study is 3,000 households from 750 DGs (125 DGs in Pure control kebeles and 625 DGs in Treatment kebeles). From each of these DGs, we randomly selected 4 households (+ 4 reserve households) to be included in the survey. Four households × 750 DGs (across 6 arms) amounts to a total sample of 3,000 households (Table A5). The baseline survey attempted to reach all 3,000 households and interviewed 2995 sample households. Primary sample households that were not available for an interview were replaced by a reserve (secondary) sample household. We will then attempt to reinterview this baseline household sample in the midline and endline survey rounds. If a household is not available during these follow-up rounds, it will no longer be replaced by a reserve household.

Table A5. Number of households by treatment arm

Arm #	Description	Number of households
1	DE on coffee	500
2	DE on coffee & maize	500
3	DE on coffee & maize + IC	500
4	DE on coffee & maize + IC + WSHG	500
5	Control	500
6	Pure control	500
	Total	3,000

¹¹ These were DGs that were listed in the second pass or had fewer than 8 coffee growing households.

Appendix B: Baseline balance

Table B1. Summary statistics of baseline characteristics, by study arm

	Full sample	1 DE on coffee	2 DE on coffee & maize	3 DE on coffee & maize + IC	4 DE on coffee & maize + IC + WSHG	5 Control	6 Pure control
Variable	Mean/(SE)	Mean/(SE)	Mean/(SE)	Mean/(SE)	Mean/(SE)	Mean/(SE)	Mean/(SE)
Household characteristics:							
Household size	5.47 (0.04)	5.49 (0.11)	5.66 (0.10)	5.38 (0.11)	5.41 (0.10)	5.38 (0.10)	5.50 (0.12)
Female headed household	0.09 (0.01)	0.10 (0.01)	0.10 (0.01)	0.09 (0.01)	0.08 (0.01)	0.11 (0.01)	0.06 (0.01)
Head's age	44.92 (0.28)	44.79 (0.67)	44.79 (0.64)	45.75 (0.77)	44.66 (0.70)	44.50 (0.66)	45.03 (0.64)
Head cannot read or write	0.43 (0.01)	0.41 (0.03)	0.45 (0.03)	0.43 (0.02)	0.42 (0.02)	0.40 (0.03)	0.46 (0.03)
Primary outcomes:							
Knowledge of good agricultural practices	5.31 (0.04)	5.32 (0.08)	5.26 (0.09)	5.30 (0.10)	5.40 (0.10)	5.22 (0.08)	5.35 (0.09)
Adoption of good agronomic practices - coffee	2.35 (0.03)	2.39 (0.06)	2.36 (0.06)	2.28 (0.08)	2.40 (0.07)	2.32 (0.06)	2.35 (0.07)
Adoption of good agronomic practices - maize	3.94 (0.03)	3.86 (0.10)	4.00 (0.07)	3.94 (0.08)	4.01 (0.07)	3.94 (0.08)	3.88 (0.08)
(log) Coffee yield	5.53 (0.02)	5.46 (0.06)	5.46 (0.05)	5.55 (0.05)	5.56 (0.06)	5.60 (0.05)	5.53 (0.06)
(log) Maize yield	7.45 (0.01)	7.45 (0.04)	7.48 (0.03)	7.39 (0.04)	7.50 (0.03)	7.44 (0.04)	7.44 (0.04)
Asset index	-0.00 (0.05)	0.12 (0.14)	0.04 (0.11)	-0.00 (0.11)	-0.10 (0.11)	-0.07 (0.11)	0.02 (0.12)
(log) Household per capita consumption expenditure	0.73 (0.01)	0.72 (0.03)	0.71 (0.03)	0.72 (0.03)	0.77 (0.03)	0.74 (0.03)	0.71 (0.03)
Secondary outcomes:							
Household per capita incomes	10.83 (0.02)	10.84 (0.05)	10.80 (0.04)	10.77 (0.04)	10.89 (0.05)	10.81 (0.05)	10.87 (0.04)
Food Insecurity Experience Scale (FIES)	3.11 (0.06)	3.03 (0.13)	2.98 (0.14)	3.39 (0.14)	3.24 (0.14)	2.94 (0.13)	3.08 (0.14)
Household Diet Diversity Score (HDDS)	7.58 (0.03)	7.57 (0.08)	7.52 (0.07)	7.54 (0.07)	7.65 (0.07)	7.51 (0.07)	7.69 (0.07)
Able to secure a loan from formal sources	0.03 (0.00)	0.03 (0.01)	0.04 (0.01)	0.04 (0.01)	0.05 (0.01)	0.03 (0.01)	0.02 (0.01)
Able to secure a loan from informal sources	0.35 (0.01)	0.34 (0.03)	0.37 (0.02)	0.32 (0.02)	0.34 (0.02)	0.34 (0.02)	0.37 (0.02)

Table B1. Balance tests of baseline characteristics**Panel A: Balance tests between study arms 1 to 5**

Variable	(1)-(2)	(1)-(3)	(1)-(4)	(1)-(5)	(2)-(3)
	Pairwise t-test Mean difference	Pairwise t-test Mean difference	Pairwise t-test Mean difference	Pairwise t-test Mean difference	Pairwise t-test Mean difference
Household characteristics:					
Household size	-0.171	0.115	0.089	0.118	0.285**
Female headed household	0.008	0.019	0.021	-0.010	0.011
Head's age	-0.003	-0.964	0.134	0.287	-0.961
Head cannot read or write	-0.043	-0.020	-0.006	0.009	0.023
Primary outcomes:					
Knowledge of good agricultural practices	0.055	0.017	-0.086	0.099	-0.038
Adoption of good agronomic practices - coffee	0.029	0.111*	-0.016	0.067	0.082
Adoption of good agronomic practices - maize	-0.144	-0.087	-0.149	-0.083	0.057
(log) Coffee yield	0.006	-0.088	-0.093*	-0.132**	-0.094
(log) Maize yield	-0.028	0.057	-0.050	0.004	0.085*
Asset index	0.081	0.126	0.222*	0.197*	0.045
(log) household per capita consumption expenditure	0.009	0.004	-0.051	-0.018	-0.006
Secondary outcomes:					
Household per capita incomes	0.045	0.076*	-0.041	0.038	0.031
Food Insecurity Experience Scale (FIES)	0.052	-0.359**	-0.205	0.088	-0.411**
Household Diet Diversity Score (HDDS)	0.049	0.029	-0.074	0.061	-0.021
Able to secure a loan from formal sources	-0.004	-0.002	-0.012	0.008	0.002
Able to secure a loan from informal sources	-0.027	0.023	0.003	-0.003	0.050*

Note: Standard errors were clustered at the kebele level. Fixed effects at the kebele were included in all estimation regressions for the F-tests. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. 1 = DE on coffee; 2 = DE on coffee & maize; 3 = DE on coffee & maize + IC; 4 = DE on coffee & maize + IC + WSHG; 5 = Control.

Panel B: Balance tests between study arms 1 to 5 (continued)

Variable	(2)-(4)	(2)-(5)	(3)-(4)	(3)-(5)	(4)-(5)
	Pairwise t- test Mean difference	Pairwise t- test Mean difference	Pairwise t- test Mean difference	Pairwise t- test Mean difference	Pairwise t- test Mean difference
Household characteristics:					
Household size	0.259**	0.288**	-0.026	0.003	0.029
Female headed household	0.013	-0.018	0.002	-0.029	-0.031
Head's age	0.136	0.290	1.098	1.251	0.154
Head cannot read or write	0.037	0.051	0.014	0.029	0.015
Primary outcomes:					
Knowledge of good agricultural practices	-0.141	0.044	-0.103	0.083	0.186**
Adoption of good agronomic practices - coffee	-0.045	0.038	-0.127*	-0.044	0.083
Adoption of good agronomic practices - maize	-0.005	0.060	-0.062	0.003	0.066
(log) Coffee yield	-0.099*	-0.139**	-0.005	-0.045	-0.040
(log) Maize yield	-0.022	0.033	-0.107**	-0.053	0.054
Asset index	0.141	0.116	0.096	0.071	-0.025
(log) household per capita consumption expenditure	-0.060*	-0.028	-0.054	-0.022	0.032
Secondary outcomes:					
Household per capita incomes	-0.087*	-0.007	-0.118**	-0.038	0.079
Food Insecurity Experience Scale (FIES)	-0.257	0.036	0.154	0.447***	0.293*
Household Diet Diversity Score (HDDS)	-0.123	0.012	-0.103	0.033	0.135
Able to secure a loan from formal sources	-0.008	0.012	-0.010	0.010	0.020
Able to secure a loan from informal sources	0.031	0.024	-0.019	-0.026	-0.007

Note: Standard errors were clustered at the kebele level. Fixed effects at the kebele were included in all estimation regressions for the F-tests. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. 1 = DE on coffee; 2 = DE on coffee & maize; 3 = DE on coffee & maize + IC; 4 = DE on coffee & maize + IC + WSHG; 5 = Control.

Panel C: Balance tests pure control vs study arms 1 to 5

Variable	Pure control vs 1 DE on coffee	Pure control vs 2 DE on coffee & maize	Pure control vs 3 DE on coffee & maize + IC	Pure control vs 4 DE on coffee & maize + IC + WSHG	Pure control vs 5 Control
	Pairwise t-test Mean difference	Pairwise t-test Mean difference	Pairwise t-test Mean difference	Pairwise t-test Mean difference	Pairwise t-test Mean difference
Household characteristics:					
Household size	-0.003	0.168	-0.118	-0.092	-0.121
Female headed household	0.041**	0.032*	0.022	0.020	0.051***
Head's age	-0.238	-0.235	0.726	-0.371	-0.525
Head cannot read or write	-0.046	-0.003	-0.026	-0.040	-0.054
Primary outcomes:					
Knowledge of good agricultural practices	-0.037	-0.092	-0.053	0.050	-0.136
Adoption of good agronomic practices - coffee	0.040	0.011	-0.071	0.056	-0.027
Adoption of good agronomic practices - maize	-0.020	0.123	0.066	0.129*	0.063
(log) Coffee yield	-0.071	-0.077	0.017	0.022	0.062
(log) Maize yield	0.005	0.033	-0.052	0.055	0.001
Asset index	0.107	0.025	-0.020	-0.116	-0.091
(log) household per capita consumption expenditure	0.008	-0.002	0.004	0.058	0.026
Secondary outcomes:					
Household per capita incomes	-0.021	-0.066	-0.098*	0.020	-0.059
Food Insecurity Experience Scale (FIES)	-0.050	-0.102	0.309	0.155	-0.138
Household Diet Diversity Score (HDDS)	-0.120	-0.169*	-0.148	-0.046	-0.181*
Able to secure a loan from formal sources	0.018*	0.022*	0.020*	0.030***	0.010
Able to secure a loan from informal sources	-0.035	-0.007	-0.057*	-0.038	-0.031

Note: Standard errors were clustered at the kebele level. Fixed effects at the woreda were included in all estimation regressions for the F-tests. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Appendix C: Woreda level results

Table C1. Knowledge of coffee best practices, by woreda

Best Practices	Description	% of households with correct knowledge				
		All	Limu Kosa	Mana	Seka Chekorsa	Shabe Sombo
Planting coffee seedling	Seedling planting hole preparation time	29.3	33.9	33.4	25.6	19.7
Weeding	No digging under tree canopy	38.5	39.2	47.7	34.3	28.0
Sucker selection	Suckers left at first & second thinning	43.4	36.7	51.4	40.1	46.9
IDM	Know at least one disease control method	68.4	56.1	81.9	69.6	67.2
Nutrition	Feeding coffee trees	69.5	61.4	81.4	69.9	64.6
Stumping method	Cutting method during stumping	76.1	66.4	84.3	77.3	78.5
Shade management	Shade cover knowledge	94.3	92.2	93.5	95.6	97.6

Table C2. Coffee best practice adoption on reference/visited plot, by woreda

	Total	Limu Kosa	Mana	Seka Chekorsa	Shabe Sombo
Observations	2995	944	851	661	539
Rejuvenation/stumping	10	7	20	6	6
Nutrition/organic fertilizer	19	10	25	31	12
Planting	51	51	55	56	42
Weeding	57	55	68	46	57
IDM	66	61	74	68	63
Shade	82	90	88	85	57

Table C3. Knowledge of maize best practices, by woreda

Best Practices	Description	% of households with recommended practice knowledge				
		All (N=2873)	Limu Kosa (N=882)	Mana (N=826)	Seka Chekorsa (N=661)	Shabe Sombo (N=504)
Ploughing	Ploughing knowledge (3 to 5 times)	95.8	97.5	98.1	96.5	87.9
Row planting	Row planting knowledge	38.1	48.0	19.5	39.8	49.4
NPS/DAP application	Application time & amount	75.6	90.5	55.1	68.1	97.6
UREA application	Application time & amount	5.5	1.1	5.7	11.3	5.0
Inter-cultivation ('shilshalo')	Inter-cultivation is not recommended	33.0	27.8	46.4	36.4	15.9
FAW control	Know at least one FAW control method	63.0	50.7	70.7	65.7	68.5
Storage	Know at least one recommended practice	50.1	33.6	57.0	54.0	62.5

Table C4. Maize best practice adoption, by woreda (% of households)

	Total		Limu Kosa		Mana		Seka Chekorsa		Shabe Sombo	
	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean
Plowed 3-4 times before sowing	2185	89	657	89	670	93	504	87	354	86
Seed rate per hectare of 25–30 kg	2185	40	657	59	670	31	504	32	354	32
Row planting	2185	93	657	97	670	98	504	94	354	72
Applied NPS/DAP	2185	86	657	95	670	96	504	91	354	46
Applied 100 to 150 kgs/ha of NPS/DAP	2185	2	657	1	670	2	504	2	354	1
Applied UREA	2185	84	657	93	670	95	504	87	354	42
Applied 150 to 250 kgs/ha of UREA	2185	0	657	0	670	0	504	0	354	0

Table C5. Green coffee beans and maize yield (kg/ha) by woreda

	Green coffee beans			Maize		
	Obs.	Mean	Median	Obs.	Mean	Median
Mana	851	260.2	218	670	2255.4	2000
Limu Kosa	942	278.4	220	653	2048.9	1800
Seka Chekorsa	661	368.9	280	504	1991.7	1600
Shabe Sombo	539	583.8	533	354	1537.6	1333
Total	2993	348.2	262	2181	2016.1	1714

Table C6. Green coffee beans and maize yield (kg/ha) by woreda

	Cherry price (birr/kg)			Jenfel price (birr/kg)			Green beans price (birr/kg)*		
	N	Mean	Median	N	Mean	Median	N	Mean	Median
Total	2049	29	30	1036	70	65	2420	143	148
Limu Kosa	757	29	30	230	76	70	794	144	150
Mana	759	31	30	205	70	70	787	153	150
Seka Chekorsa	246	28	28	298	73	70	411	142	139
Shabe Sombo	287	24	25	303	61	60	428	122	120

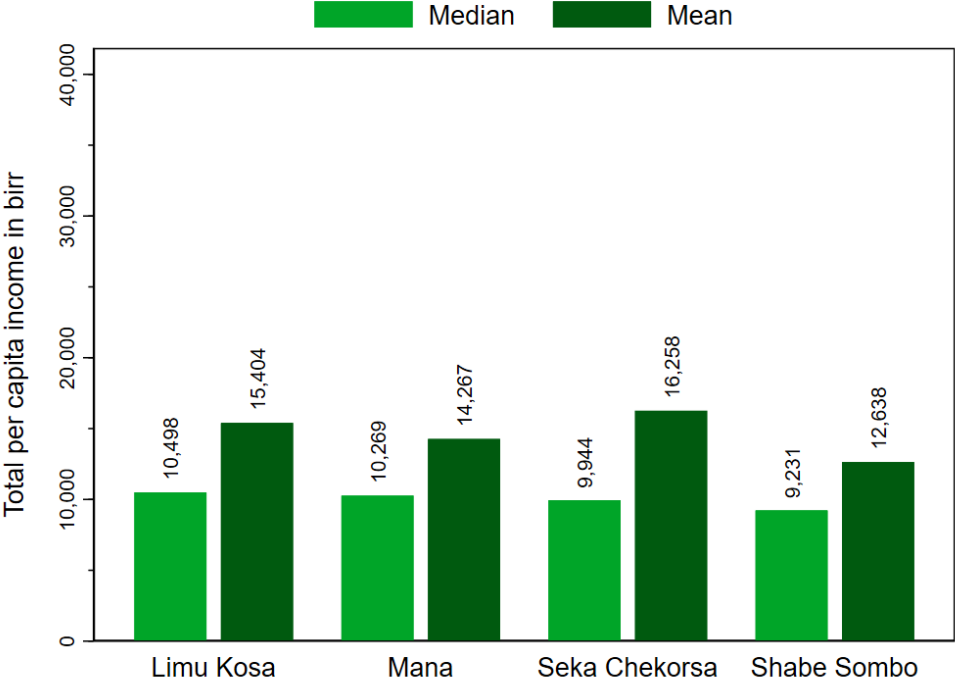
Note: *The green beans price was calculated based on the cherry and jenfel prices following the standard conversion factor from the Ethiopia Coffee Atlas book (Coffee Atlas of Ethiopia, Davis, Kebebew Challa, Williams (uchicago.edu), 2018).

Table C7. Coffee seedling planting and source, by woreda

	Total		Limu Kosa		Mana		Seka Chekorsa		Shabe Sombo	
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
Household planted coffee seedling (%)	2995	53.5	944	52.1	851	56.4	661	56.3	539	48.1
Number of coffee seedlings planted	1603	120.0	492	152.9	480	86.9	372	90.9	259	160.7
Seedling source (%)										
Own farm	2144	66.0	660	53.0	650	70.2	471	66.7	363	81.3
Private nurseries	2144	19.5	660	31.5	650	17.2	471	9.3	363	14.6
Gov't/bureau of agri./DA nurseries	2144	10.0	660	9.7	650	9.7	471	16.4	363	2.8
Cooperative nurseries	2144	3.5	660	3.0	650	2.6	471	7.0	363	1.1
Research institute	2144	0.4	660	0.6	650	0.2	471	0.4	363	0.3
Other, specify	2144	0.8	660	2.1	650	0.2	471	0.2	363	0.0

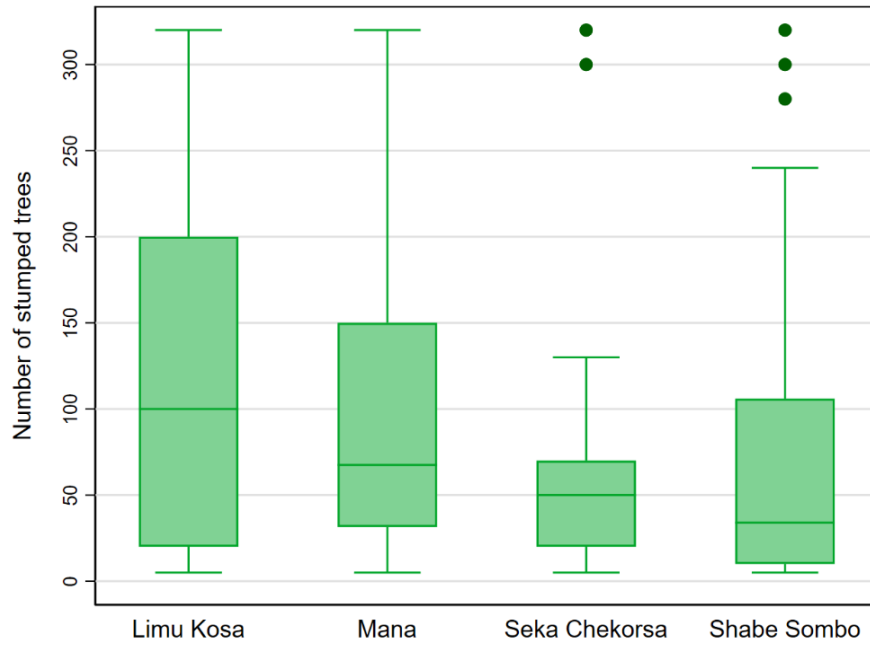
Appendix D: Supplementary figures on household welfare outcomes

Figure D1. Mean and median total per capita incomes, by woreda



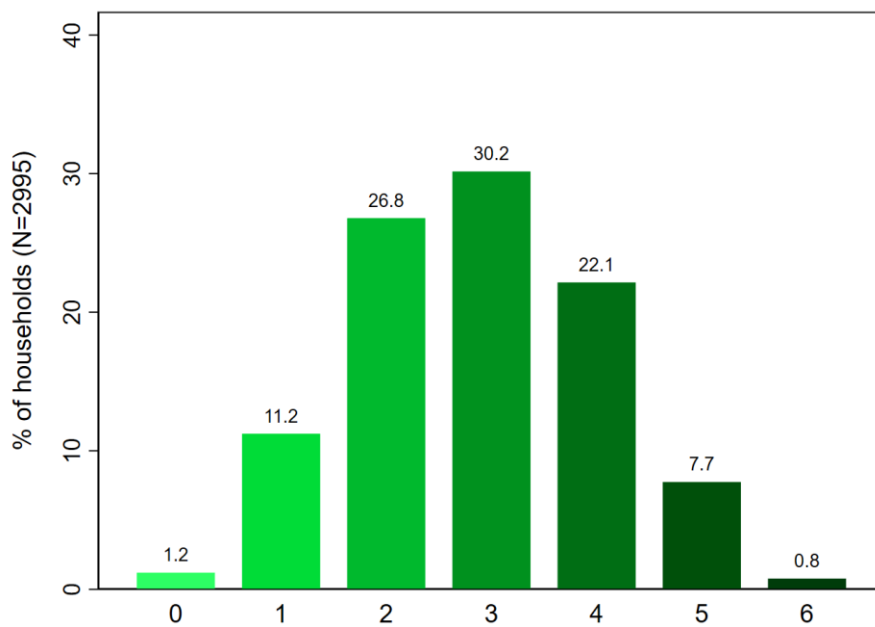
Appendix E: Supplementary figures on adoption of coffee best practices

Figure E1. Total number coffee trees stumped during the last three season (2021 – 2023), by woreda



Note: The values were winsorized at 5 and 95 percentiles to reduce the effect of outliers.

Figure E2. Number of recommended coffee practices adopted by the household



Appendix F: Supplementary figure on locations of coffee sold

Figure F1. Location where cherry coffee sold

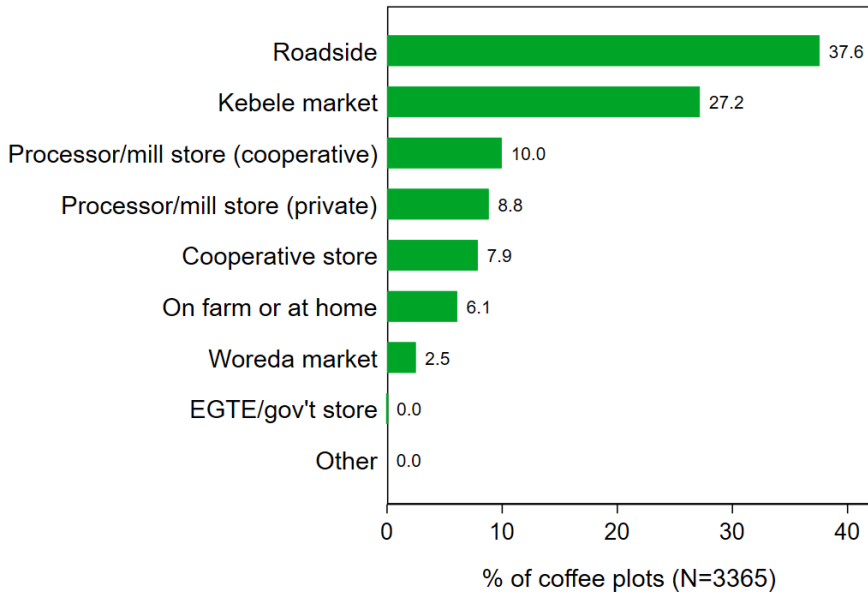
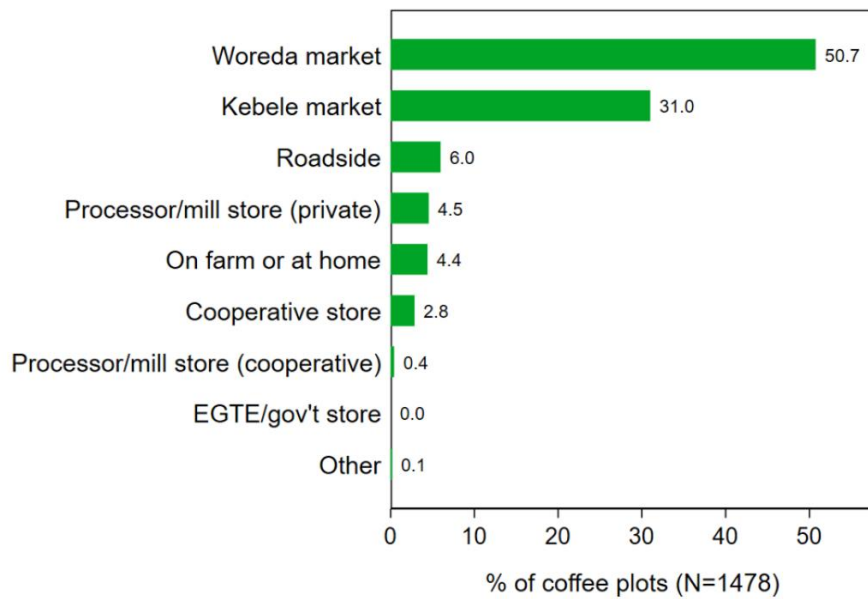


Figure F2. Location where jenfel coffee sold



Appendix G: Supplementary table on maize and coffee knowledge and adoption rules

Table G1. Maize knowledge and adoption rules

Technology	Knowledge		Adoption (on at least one of the maize plots)	
	Indicators	Rule	Indicator	Rule
Land preparation	<ol style="list-style-type: none"> Recommended frequency of ploughing: 3 to 4 times or 2 times if ploughed using tractor Leave the crop residue 	1 if provide correct for both questions, 0 otherwise	<ol style="list-style-type: none"> Recommended frequency of ploughing: 3 to 4 times Leave the crop residue: at least 30% of farmland cover 	1 if applied the recommended frequency & crop residue, 0 otherwise
Row planting	<ol style="list-style-type: none"> Recommended distance between rows: 75 to 80 cms Recommended distance between plants: 25 to 40 cms Seed depth: 5 to 12 cms Recommended seeding rate for row planting: 25 to 30 kgs/ha 	1 if correctly answer at least two questions, 0 otherwise	<p>Must be planted in rows (f95) with:</p> <ol style="list-style-type: none"> Distance b/n rows (f96): 75 to 80 cms Distance b/ plants (f97): 25 to 40 cms Seed depth (f97a): 5 to 7 cms Recommended seeding rate for row planting (f93): 25 to 30 kgs/ha 	1 if planted in rows and satisfies at least 2 of the 4 criteria, 0 otherwise
NPS/DAP application	<ol style="list-style-type: none"> NPS/DAP application time: should apply only during planting NPS/DAP amount: recommended application rate on a hectare: 100 to 150 kgs/ha 	1 if provide correct answers for both questions, 0 otherwise	<p>PS/application time (f100): apply only during planting</p> <p>NPS/DAP amount: recommended application rate: 100 to 150 kgs/ha</p>	1 if applied the recommended amount at planting, 0 otherwise
UREA application	<p>UREA application time:</p> <ol style="list-style-type: none"> First round: 20 to 25 days after planting Second round: 35 to 40 days after planting UREA amount: recommended application rate on a hectare: 100 to 150 kgs/ha 	1 if correctly answered the recommended rate AND at least the first OR second application time, 0 otherwise	<ol style="list-style-type: none"> UREA amount: recommended application rate: 150 to 250 kgs/ha UREA application time: First round: 20 to 25 days after planting Second round: 35 to 40 days after planting 	1 if applied the recommended rate of UREA; AND the first OR the second round applied at the recommended time, 0 otherwise
Inter-cultivation ('shilshalo')	Shilshalo/Inter-cultivation is not a recommended practice	1 if provide correct, 0 otherwise	Shilshalo/Inter-cultivation is not a recommended practice (fm7)	1 if not applied, 0 otherwise
Fall Armyworm (FAW) control	List at least one FAW control method	1 if provide correct answer, 0 otherwise	<ol style="list-style-type: none"> FAW incidence in the last 3 year (fm9) Cultural FAW control methods (fm10) Maize plot inspection to check for pests & disease (fm8) 	1 if there was FAW incidence & at least one control method used. OR 1 if there was no FAW incidence but the farmer inspected the maize plot weekly or more frequently
Storage	<p>Recommended storage practices:</p> <ol style="list-style-type: none"> Dry grains before storage; 2. clean storage; 3. Dry storage; 4. Fully sealed storage 5. Treat grain if it will be stored for more than 3 months 	1 if know at least three recommended practices listed, 0 otherwise	<p>Recommended storage (fm14):</p> <ul style="list-style-type: none"> Modern storage (metal silos) Normal sacks PICs bag (bag treated with chemical) 	1 if the farmer stored the maize grains on one of the recommended storages, 0 otherwise

Table G2. Coffee knowledge and adoption rules

Technology	Knowledge		Adoption	
	Indicators	Rule	Indicator	Rule
Weeding	Is it a good practice to weed by digging up the tree canopy?	1 if the farmer responded “false”, 0 otherwise	1. Weeding Frequency 2. Weed status 3. Weed Size 4. Weeding Method	Adopted if farmer: 1. Weeded twice or more per year 2. Has no weeds 3. Farmer has not dug under the canopy OR 1. Weeds twice or more per year 2. Has few weeds under the canopy 3. Weeds are less than 30cm tall 4. Farmer has not dug under the canopy
Nutrition (Organic fertilizer)	How can you feed you coffee?	1 if the farmer mentioned fertilizer, 0 otherwise	Organic fertilizer uses in the past 12 months	Adopter if applied any type of organic fertilizer on the best practice plot, 0 other wise
Stumping	1. Recommended cutting/stumping method: slash or 45° 2. Height where the stem to cut/stumped: 30 cms from the ground	1 if the farmer provided at least one correct answer, 0 otherwise	Stumping of coffee trees on the best practice plot in the past three year	Adopter if farmer had stumped during any of the past three years, 0 otherwise
Sucker selection	1. Suckers should keep in the first round of thinning: 3 to 4 2. Suckers should keep in the first round of thinning: 2 to 3	1 if the farmer provided at least one correct answer, 0 otherwise	-	-
Planting	When should one prepare the coffee seedling planting hole/pit?	1 if the farmer provided at least one correct answer, 0 otherwise	1. Planting of at least 20 new coffee trees in the past 1 year 2. Own coffee nursery in the past 1 year	Adopter if a farmer planted at least 20 trees on the best practice plot; OR Owned a coffee nursery, 0 otherwise
Integrated disease management (IDM)	Methods known to reduce the incidence of Coffee Berry Disease (CBD)	1 if the farmer mentioned at least one CBD control method, 0 otherwise	1. Main disease problem in the past 1 year 2. Disease (CBD) control method 3. Coffee farm inspection	Adopter if there was disease incidence & the farmer applied at least one control method; OR There was no incidence, but the farmer inspected the coffee farm weekly or more frequently, 0 otherwise
Shade management	Is it good practice to grow your coffee in the full sun, with no shade?	1 if the farmer responded “no”, 0 otherwise	1. Shade coverage 2. Planting new shade trees in the past 2 years	Adopter if there was 20% or more; OR There was less than 20% shade but shade trees planted in the last 2 years, 0 otherwise