

---

# LIVELIHOODS IN ETHIOPIA

Impact evaluation of linking smallholder coffee  
producers to sustainable markets

---

Effectiveness Review Series

2014/15

---



Photo credit: Sven Torfinn/Oxfam

ANDREW ANGUKO

OXFAM GB



# CONTENTS

Acknowledgements .....	3
Executive Summary.....	4
1 Introduction.....	9
2 Project description .....	12
3 Evaluation design .....	13
4 Data collection .....	14
4.1 Sampling approach.....	14
4.2 Analysis .....	16
5 Results.....	17
5.1 Introduction.....	17
5.2 Involvement in project activities .....	18
5.3 Quantity of coffee produced and revenues from sales .....	19
5.4 Perceptions of changes in: Coffee production, income, prices and processing machines .....	21
5.5 Increased use of modern agricultural practices .....	22
5.6 Agricultural production .....	24
5.7 Quantity of sales from agricultural production .....	25
5.8 Revenue from agricultural sales .....	26
5.9 Overall household income .....	27
5.10 Household asset wealth.....	29
5.11 Care module .....	30
5.11.1 Time use on care.....	31
5.11.2 Hours spent in a day on care and productive activities .....	31
5.11.3 Redistribution of responsibilities for care within the household .....	33
5.11.4 Ownership of time/labour-saving equipment.....	36
5.11.5 Attitude towards and awareness of care work.....	37
6 Conclusions .....	38
6.1 Conclusions .....	38
6.2 Programme learning considerations .....	39
Appendix 1: Baseline statistics before matching .....	41
Appendix 2: Methodology used for propensity-score matching .....	42
Appendix 3: Robustness checks .....	47
Notes .....	52

## ACKNOWLEDGEMENTS

We would like to thank Oxfam Ethiopia and local partner Limmu Inara Cooperative Union, particularly Fekadu Dugasa, the union manager, for being so supportive during this exercise. Particular thanks are due to Rahel Bekele, Project Manager and Mulu Tesfaye, Livelihood Programmes Coordinator, as well as Bisrat Markos, the consultant who managed the survey work.

# EXECUTIVE SUMMARY

Oxfam GB's Global Performance Framework is part of the organisation's effort to better understand and communicate its effectiveness, as well as enhance learning across the organisation. Under this Framework, a small number of completed or mature projects are selected at random each year for an evaluation of their impact, known as an 'Effectiveness Review'. The project 'Coffee Value Chain – Linking Smallholders to a Sustainable and Scalable Business Model in Ethiopia' (ETHB29) was one of those selected for an Effectiveness Review in the 2014/15 financial year.

The project activities were implemented by Oxfam GB in conjunction with Limmu Inara Multipurpose Cooperative Union. The project was started in September 2009 and was completed in March 2012. The project's overall objective was to contribute to **improved Coffee production and sales** by linking coffee farmers to primary coffee cooperatives and unions. The union in this case was Limmu Inara Multipurpose Cooperative Union. The project covered three districts: Limmu Seka, Limmu Kosa and Chora Botter where there were primary coffee cooperatives that bought coffee directly from the farmers and then delivered the coffee to Limmu Inara Cooperative Union. The union supplied coffee seeds, slashers and wire mesh and supported the farmers in coffee seedling production. Training sessions were organised after a capacity assessment to identify the gaps among the coffee-growing farmers in the region. The capacity building training sessions were organised in collaboration with Jimma University in Ethiopia. Oxfam also identified an organisation called Farm Organic International (FOI), which promoted the marketing of coffee internationally on behalf of the union. The organisation specifically identified markets for coffee and carried out capacity building for the union board. In turn, Oxfam provided funds for project implementation and paid for the services of Farm Organic International. In addition, Oromiyaa Cooperative Bank provided credit facilities to the coffee farmers through the cooperative societies.

The Functional Adult Literacy (FAL) scheme spearheaded by Limmu Inara Cooperative Union was aimed at building the capacity of coffee farmers, particularly in the use of modern agricultural practices with a view to improving the quantity and quality of coffee produced.

## EVALUATION APPROACH

The review adopted a quasi-experimental impact evaluation design, which involved comparing households that had been supported by the project with households in neighbouring communities that had not been supported by the project, but had similar characteristics in 2009 before the project was implemented.

The Effectiveness Review was carried out in the districts where the project activities had been implemented since at least 2009. Within those districts, the households that had participated in the project during the period were selected at random to be interviewed. For comparison purposes, interviews were also carried out with farmer households that had not participated in the project, but who were eligible and had expressed an interest in doing so. In total, 215 project participants and 432 non-participants were interviewed. At the analysis stage, the statistical tools of propensity-score matching and multivariate regression were used to control for demographic and baseline differences between the households surveyed in project and comparison areas, to provide confidence when making estimates of the project's impact.

# RESULTS

The results from this Effectiveness Review suggest that the project under analysis has had a positive effect on households that participated in the project on variety of outcomes.

One key question for this review was to check whether the quantity of coffee produced had changed. Survey respondents were asked whether any member of their household cultivated coffee in the past 12 months. If yes, they were asked to state the quantity of coffee the household harvested in the last 12 months. Results indicate that on average, households that participated in the project harvested 64.7 per cent more coffee than non-participants.

A second question was to determine whether the quantity of coffee sold had changed. Respondents were asked to state the quantity of coffee the household sold in the past 12 months. The study found that on average, households that participated in the project sold 1.28 times, or 128 per cent, more coffee than their counterparts in comparison areas.

Thirdly, respondents were asked to approximate the total value of sales of coffee produced in the past 12 months. Results indicate that on average, households in intervention areas obtained 1.16 times, or 116 per cent, more revenue than non-participants. Further analysis revealed that the total quantity of all crops (coffee included) produced, sold, and the revenues obtained, was significantly higher in intervention areas than in the comparison.

The project participants received agricultural input, such as seeds, fertilisers, wire mesh, jute sacks and capacity-building training organised by the Cooperative Union. They were also supported in coffee seedling production. The Cooperative Union increased accessibility to coffee markets through the primary coffee cooperative societies. These activities led to an increase in coffee production, sales and revenues in the intervention areas.

The Effectiveness Review also sought to find out whether coffee farmers adopted modern agricultural practices acquired through capacity building training offered by the Cooperative Union. The agricultural practices included the use of improved seedlings, sun dryers, compost manure and organic fertilisers. Results indicate that on average, there was a 3.4 percentage point increase in the households in intervention areas that adopted any of the modern agricultural practices, compared with households in comparison areas. The Effectiveness Review also revealed a 5.2 percentage point increase in participant households adopting use of compost manure and other organic fertilisers compared with non-participants.

It was also important to investigate self-reported perceptions regarding changes in coffee production, income from coffee sales, prices, and the number of coffee processing machines. On average, an 18.9 percentage point increase in households in project areas reported increased coffee production compared with non-project participants. Similarly, about 17.3 percentage point and 10.1 percentage point increase in households in intervention areas reported increased income and prices of coffee respectively compared with non project participants. In addition, there was a 23.7 percentage point increase in the number of participant households reporting an increase in the number of coffee processing machines compared with non-participants.

In a separate section of the survey, respondents were asked for details about their household's recent consumption, of both food and non-food items. An overall figure for per-day, per-person consumption was calculated, to serve as a proxy measure for net

household income. Consumption was estimated to be approximately 13.6 per cent higher among project participants than non-participants, suggesting that the project has had a corresponding effect on overall household income. The overall household income was 28 per cent higher on average, among participant households than non-participants and was statistically significant different from zero. This indicator also represents the new Oxfam GB Global Indicator for Livelihoods.

This conclusion is reinforced by examining changes in indicators of households' material wealth (ownership of assets and housing conditions) between 2009 and the date of the survey. Project participants' material wealth has increased in terms of wealth indicators since 2009. The evidence that the project participants have significantly increased their household income adds weight to the conclusion that the coffee cooperative unions are providing a valuable marketing channel to coffee producers, and that project participants have generally been receiving higher prices than they otherwise would have.

#### Key results of this Effectiveness Review

Outcome	Evidence of positive impact	Comments
Increased quantity of coffee produced	YES	On average intervention households produced 64.7 per cent more coffee than households in comparison communities.
Increased quantity of coffee sold	YES	Participant households sold 1.28 times or 128 per cent more coffee, on average, than households in comparison communities.
Increased revenue from coffee sales	YES	Intervention households obtained, on average, 1.16 times or 116 per cent more revenue from coffee sales than comparison communities.
Household Asset Wealth (Wealth Index)	YES	Households in intervention areas are more likely to have higher household asset wealth than households in comparison communities.
Overall Household income (New Global Indicator)	YES	Households in project areas had a 28.01 per cent increase in overall household income compared with households in the comparison communities. This indicator represents the new Oxfam GB Global Indicator for livelihoods and was statistically significant different from zero.
Increased use of modern agricultural practices (any of the practices)	YES	On average, there was a 3.4 percentage point increase in the number of participant households reporting increased use of modern agricultural practices compared with non-participants.
Increased practise of compost manure and organic fertilizers	YES	On average, there was a 5.2 percentage point increase in the number of participant households reporting increased use of compost manure and organic fertilisers compared with non-participants.

In order to provide some continuity with previous Effectiveness Reviews, the old Global Livelihoods Indicator has also been reported in this Effectiveness Review (Table 5.8). It assesses whether or not the income of each household has moved above an arbitrary line (defined by the median household income in the comparison group). The indicator captures increases in income that cross the threshold (the median income among the comparison group). A very poor household (in the lower half of the income distribution) may have increased its income substantially, but this is not relevant for the global indicator unless they reach the threshold. Similarly, a household that would have been

in the top half of the income distribution (i.e. with above-average income) would never be counted towards the indicator. Only households that started below the threshold and experienced an increase in income to take them above the threshold (median in the comparison) are counted positively. In this Effectiveness Review, there was a 21.2 percentage points change in average household income (as measured by household consumption and expenditure per adult equivalent person, per day)

It is important to note that all the estimates provided here are robust to other estimation models presented in Appendix A3.

## PROGRAMME LEARNING CONSIDERATIONS

### **Strengthening marketing channels in livelihood programmes leads to better incomes for farmers.**

The results of this Effectiveness Review add to the proof that farmers' markets provided by the coffee cooperative unions have a positive impact in terms of increasing income. This Effectiveness Review provides clear evidence that the new marketing channels for coffee farmers (selling through the cooperatives), which cut out intermediaries, and the related support provided to producers under this project, have led to a significant improvement in household income. This evidence can be used to strengthen the case both for scaling up the farmers' markets in Ethiopia and for encouraging adoption in other parts of Ethiopia where coffee is grown.

### **Engaging stakeholders with clearly defined roles should be considered.**

The critical role of the stakeholders in moving the project success forward through engaging them from the outset, contributed to positive outcomes. One key learning from this Effectiveness Review is the pivotal role played by the bureaux and sectoral institutions, such as Jimma University and the cooperatives agency, in delivering an integrated support for success, influencing by demonstration and scaling. Jimma University led in capacity-building training to equip coffee farmers with the necessary skills for coffee production. The cooperatives unions facilitated the marketing of coffee and ensured that the farmers sold greater quantities and obtained higher revenues for their produce.

### **Capacity building of beneficiaries both before and during project implementation should be encouraged in livelihoods interventions.**

The Functional Adult Literacy (FAL) scheme spearheaded by Limmu Inara Cooperative Union was aimed at building the capacity of coffee farmers, particularly in the use of modern agricultural practices, with a view to improving the quantity and quality of coffee produced. Indeed, there is evidence from this Effectiveness Review that farmers adopted modern methods of farming. The Integrated Functional Adult Learning contributed to the acquisition of new skills and knowledge and should therefore be encouraged in future implementation strategies in projects of a similar nature. Leveraging on external stakeholders and global events for demonstration, synergy and influencing, such as the global PS learning events organised by Oxfam, Catholic Relief Services (CRS), Slow Food and a finance fare with BioFach should also be encouraged in implementing future projects.

**The identification of stages within a given value-chain where women can actively engage should be considered.**

One important learning consideration is that when designing a project, it is important to consider in which stages of the value chain women best fit and what kind of special support they may require. At times, it may not matter whether the commodity is a cash crop mainly dominated by men. However, knowing which area works best for women is crucial. In this particular coffee value chain project, women engaged in coffee seedling preparation, and this proved to be quite relevant for them and consequently more seedlings were produced, which ultimately boosted coffee production.

**Institutional strengthening of the Limmu Inara Cooperative Union worked better than support through a business service provider.**

Initially, Oxfam commissioned a business service provider to support the Limmu Inara Cooperative Union to play a facilitation role and link them with the export buyers. One of the lessons learnt is that strengthening the institutional capacity of the union can make it more effective in discharging its functions. The union hired its own commercial manager to facilitate its own export link. As a result, the union managed to directly connect to export buyers and was able to negotiate for prices. This ensured that the union had more potential buyers who could offer better coffee prices. Consequently, coffee farmers were able to sell more of their produce at better prices and hence obtain higher revenues.

# 1 INTRODUCTION

Oxfam GB has put in place a Global Performance Framework (GPF) as part of its effort to better understand and communicate its effectiveness, as well as to enhance learning across the organisation. This framework requires project and programme teams to annually report output data across six thematic indicator areas. In addition, every year, for each thematic indicator area a modest sample of mature<sup>1</sup> projects are randomly selected to be evaluated through rigorous Effectiveness Reviews. One key focus is on the extent that they have promoted change in relation to relevant OGB global outcome indicators.

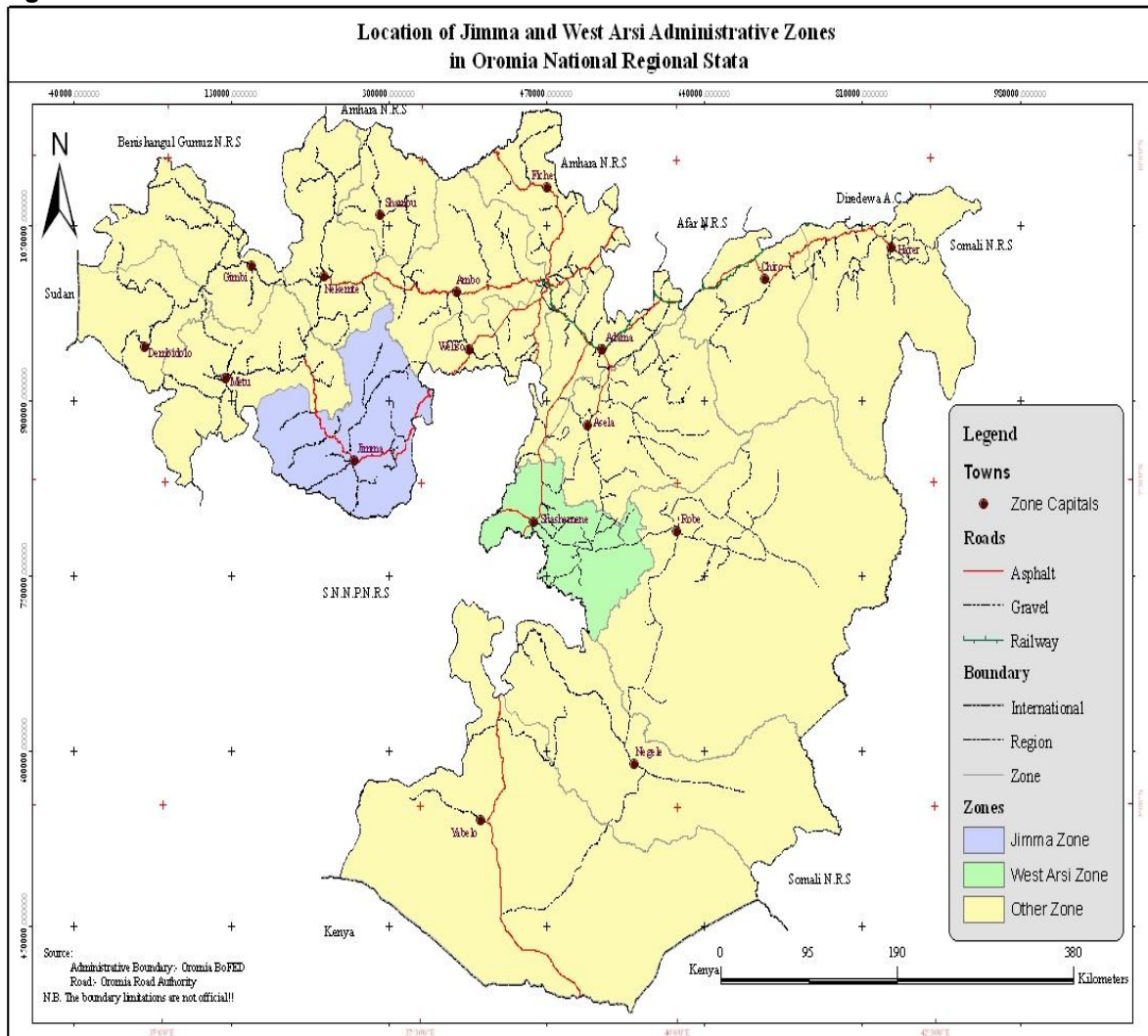
The global outcome indicator for the livelihoods thematic area is defined as the percentage of households demonstrating an increase in household income as measured by household expenditure per capita, compared to a 'typical' comparison household. This indicator is explained in more detail in Section 5 below.

The Effectiveness Review that took place in November/December 2014 was intended to evaluate the success of the project 'Coffee Value Chain – Linking Smallholders to Sustainable Markets in Ethiopia' in promoting food security and strengthened coffee production and marketing among the project participants. This project was implemented in three districts in Oromiyaa regional state – Limmu Kossa, Limmu Seka and Chora Botter – between March 2009 and April 2012 by Oxfam in conjunction with Limmu Inara Multipurpose Cooperative Farmers' Union. There were 10 primary cooperatives in Limmu Seka and seven primary cooperatives in Limmu Kosa with only one primary cooperative in Chora Botter.

The focus of the review was on the three districts supported by the project from 2009 through to its completion in April 2012. This was a three-year project targeting the three districts in Oromiyaa regional state. Most of the activities were implemented in two districts – Limmu Seka and Limmu Kosa – with less intensity of activities in Chora Botter. Consequently, some villages in Chora Botter and other nearby districts acted as comparison group. The activities implemented included Functional Adult Literacy (Adult Learning Capacity Building) among coffee farmers in the target districts, improving grinding mills, linkage to markets, inputs (such as jute bags, wire mesh, slashers, seeds and organic fertilisers, coffee seedling production) and credit facilities to the coffee farmers.

This report presents the findings of the Effectiveness Review. Section 2 briefly reviews the project description. Section 3 describes the evaluation design used, and Section 4 describes how this design was implemented. Section 5 presents the results of the data analysis, based on the comparison of outcome measures between the intervention and comparison groups. Section 6 concludes the document with a summary of the findings and some programme learning considerations. Finally, baseline statistics before matching are provided in Appendix 1 while technical and methodological considerations on the propensity-score matching (PSM) are given in Appendix 2.

**Figure 2.1: Location of intervention districts**



Source: Baseline survey of project – Linking Smallholder Coffee Producers to Sustainable Markets in Ethiopia



## 2 PROJECT DESCRIPTION

Oxfam has embarked on a long-term agriculture scale-up programme to contribute to poverty reduction efforts in Ethiopia through reaching smallholder farmers using a market-based value-chain approach.

The programme focused on by this evaluation is a coffee-based value-chain project that operates in collaboration with Limmu Innara Multipurpose Cooperative Union. The union was established in 2006 and OGB started a pilot project in 2007 and launched a three-year coffee-value-chains support project in 2009. At present the union embraces 49 primary cooperatives in three districts of Jima Zone: Limmu Seka, Limmu Kossa and Chora Botter, of which 41 are coffee producers and 32 are actively working with the union.

The project integrates Oxfam's Regional Innovation Fund support of Functional Adult Learning within value chains for better community livelihoods and economic empowerment. The programme, which focused on developing a sustainable and scalable business model, intended to benefit 13,000 smallholder coffee farmers engaged in organic coffee production. It also focused on developing a market for organic coffee that has the potential to contribute significantly towards the improvement of smallholder farmers' livelihood opportunities, especially women.

The overall objective of the project was to contribute to strengthening the socio-productive process of sustainable development by promoting coffee production and marketing initiatives in three districts in the Oromiyaa regional state of Ethiopia (Figure 2.1). Specifically, the project aimed at strengthening the productive capacity of 3,000 households by improving their production systems, ensuring sustainable food security, increasing coffee production, sales and access to coffee markets, and improving coffee prices.

The activities involved provision of quality coffee seeds, support in coffee seedling production and support in training, coffee handling, drying washing and credit facilities to farmers, and ensuring that the farmers sold their coffee to the primary cooperative unions, which in turn sold coffee to Limmu Innara Cooperative Union. The project was also involved in the provision of mesh wire, jute sacks, coffee seedlings and slashers. The programme aimed to increase household income and food security through improvement in coffee production, sales and revenues.

Higher coffee production and market linkages for increased coffee sales was expected to lead to better revenues and hence increased income for the households involved in the project. At the same time, membership of primary coffee cooperative societies was intended to ensure that farmers have access to markets and credit facilities, which will enable the purchase of agricultural inputs with the ultimate aim of increasing coffee production and sales.

# 3 EVALUATION DESIGN

The central problem in evaluating the impact of any project is how to compare the outcomes that result from that project with *what would have been the case* without that project having been carried out. In the case of this Effectiveness Review, information about the lives and livelihoods of project participants was collected through a household questionnaire – but clearly it was not possible to observe what their situation would have been had they not had the opportunity to participate in this project. In any evaluation, that ‘counterfactual’ situation cannot be directly observed: it can only be estimated.

In the evaluation of programmes that involve a large number of units (whether individuals, households, or communities), common practice is to make a comparison between units that were subject to the programme and those that were not. As long as the two groups can be assumed to be similar in all respects except for the implementation of the specific project, observing the situation of those where the project was not implemented can provide a good estimate of the counterfactual.

An ideal approach to an evaluation such as this is to select at random the areas in which the project will be implemented. Random selection minimises the probability of there being systematic differences between the project participants and non-participants, and so maximises the confidence that any differences in outcomes are due to the effects of the project.

In the case of the project examined in this Effectiveness Review, the unit at which the programme was implemented was the household within selected coffee cooperative societies. The project was implemented within the 15 primary coffee cooperative societies that were involved in the project. The selection of cooperatives involved in the project was not made at random; in fact, cooperatives were deliberately chosen based on their being particularly vulnerable in terms of low coffee production among the members, low quality coffee, low revenues and lack of access to coffee markets. The project worked with coffee farmers that were members of the primary cooperative societies. However, discussions with the implementation staff revealed that there were in fact other coffee cooperatives that were considered also suitable for implementation than could, in fact, have been covered by the project.

There were neighbouring areas with many coffee cooperatives with members who would be eligible and had the potential to benefit from participation, but who did not participate. These cooperatives had not started selling coffee to the Limmu Inara Multipurpose Cooperative Union because they had not so far learned about this opportunity. This implied that it was possible to identify non-participants who were members of these cooperatives and who are not systematically different from those project participants, who could function as a comparison group.

This allowed a ‘quasi-experimental’ evaluation approach to be adopted, in which the situation of non-participant coffee farmers was assumed to provide a reasonable counterfactual for the situation of farmers who had participated in the project activities.

To improve the confidence in making this comparison, households in the project coffee cooperatives were ‘matched’ with households with similar characteristics in the non-project (or ‘comparison’) coffee cooperatives. Matching was performed on the basis of a variety of characteristics – including household size, education level, productive activities and indicators of material wellbeing, such as housing conditions and

ownership of assets. Since some of these characteristics may have been affected by the project itself (particularly those relating to productive activities and wealth indicators), matching was performed on the basis of these indicators *before* the implementation of the project. Although baseline data was available, it did not have all the variables that could be used during matching and so survey respondents were asked to recall some basic information about their household's situation from 2009, before the project was implemented. While this recall data is unlikely to be completely accurate, this should not lead to significant bias in the estimates as long as the measurement errors due to the recall data are not significantly different for the treatment and comparison group.

Recall survey data provided a variety of baseline household characteristics on which matching could be carried out. These characteristics were used to calculate a 'propensity score', which is the conditional probability of the household being a participant, given the set of observable characteristic in the baseline. Project households and comparison households were then matched based on their having propensity scores within certain ranges. Please refer to Appendix 2 for a more extensive explanation of the matching procedure and tests carried out after matching to assess whether baseline characteristics are similar between the two groups.

As a check on the results derived from the propensity-score matching process, results were also estimated using multivariate regression models. Like propensity-score matching, multivariate regression also controls for measured differences between intervention and comparison groups, but it does so by isolating the variation in the outcome variable explained by being a project participant, after the effects of other explanatory variables have been accounted for.

It should be noted that both propensity-score matching and multivariate regression rely on the assumption that the 'observed' characteristics (those that are collected in the survey and controlled for in the analysis) capture all of the relevant differences between the two groups. If there are 'unobserved' differences between the groups that matter for project participation, then estimates of outcomes derived from them may be misleading. Unobserved differences between the groups could potentially include differences in attitudes or motivation (particularly important when individuals have taken the initiative to participate in a project), differences in community leadership, or local-level differences in weather or other contextual conditions faced by households. The choice of which intervention and comparison cooperatives to survey for this Effectiveness Review was made principally to minimise the potential for any such unobservable differences to bias the results.

## 4 DATA COLLECTION

### 4.1 SAMPLING APPROACH

The intervention group consisted of coffee farmer households that participated in the project from inception to closure. The intervention group was selected based on its vulnerability. Since it was not possible to include all coffee cooperatives/households due to insufficient resources, some cooperatives/households did not take part in the project activities even though they were also vulnerable. This group therefore formed the comparison group. It is important to take note of these two terms, *intervention* and *comparison*, since they are used frequently in this report.

Cooperative societies were randomly selected with probability proportional to size. This meant that societies that had more membership of coffee farmers were more likely to be selected. Using this method, seven primary coffee cooperative societies were selected from a list of 12 cooperatives that participated in the project. After the selection of the cooperative societies, lists of coffee farmers were obtained from the selected primary cooperative societies. The number of farmers to be selected from each of the seven cooperative societies was determined by dividing the number of members in each of the selected cooperatives by the total membership in the seven cooperatives and multiplying by the sample size that had been determined for the intervention group. The result provided the number of coffee farmers to be interviewed from each of the seven selected cooperative societies as indicated in column 5 of Table 4.1.

Table 4.1 shows the cooperatives and number of households sampled from the intervention and comparison group. The sample frame was constructed by identifying those households that received at least one project intervention line from the project area. Among the households that benefited directly from the project, 215 were randomly selected to be interviewed.

Lists of members of the selected cooperative societies were obtained from the chairmen of the societies. Using these lists, systematic random sampling was used to identify the coffee farmers to be interviewed. In the comparison areas, five primary cooperative societies were randomly sampled from a list of 11 primary coffee cooperatives that did not participate in the project. Lists of members were obtained from the sampled cooperatives and then systematic sampling was used to identify the farmers to be interviewed. A total of 432 households were sampled in the comparison group.

A household questionnaire was developed by Oxfam staff, in collaboration with partners, to capture data on various outcome and intervention exposure measures associated with the project's activities. Demographic data and recalled baseline data were also collected to statistically control for differences between the supported and comparison households that could not plausibly be affected by the project. The questionnaire was pre-tested by Oxfam local staff and then by the enumerators during a practice exercise, and revised accordingly.

A team of 12 enumerators (four men and eight women) was locally recruited from Limmu Seka and Limmu Kosa districts of Oromiyaa state in Ethiopia. These enumerators participated in a two-day training workshop, which was led by Oxfam staff. The second day of the workshop involved a piloting exercise, where a community in Jimma was identified. Following this exercise, the performance of each of the enumerators was reviewed individually before their appointments were confirmed.

The enumerator team was divided into two groups and mobilised to the villages represented by the seven cooperative societies in the intervention areas and five cooperatives in the comparison areas. The movement plan was created in consultation with the overall field supervisor to ensure that completed surveys were collected and reviewed at the end of every day. Feedback was provided regularly to all enumerators regarding their performance.

The full list of kebeles (sub-districts) with numbers of coffee cooperatives included in the Effectiveness Review, and the numbers of households interviewed in each, is shown in Table 4.1.

**Table 4.1: Intervention and comparison groups sample sizes**

	Project participants		Sample project participants		Sample comparison group	
	<i>Cooperatives participating in the project</i>	<i>Households/farmers participating in the project</i>	<i>Cooperatives surveyed</i>	<i>Households/farmers interviewed</i>	<i>Cooperatives selected in comparison communities</i>	<i>Households/farmers interviewed in comparison communities</i>
Bufata Gibe	1	236	1	16	0	0
Debello	1	718	1	51	0	0
Jato Seka	1	440	1	31	0	0
Kema Kube	1	865	1	61	0	0
Kiltu Cheba	1	360	1	26	0	0
Shogale	1	138	1	10	0	0
Tencho	1	315	1	20	0	0
Siltu	0	0	0	0	1	184
Laku Cime	0	0	0	0	1	37
Karayu Kobi	0	0	0	0	1	82
Sabi Warabo	0	0	0	0	1	72
Walataenso Cheko	0	0	0	0	1	57
<b>TOTAL</b>	<b>7</b>	<b>3072</b>	<b>7</b>	<b>215</b>	<b>5</b>	<b>432</b>

## 4.2 ANALYSIS

Households of project participants and non-participants were compared in terms of their demographic characteristics, livelihoods activities and economic situation in 2009. These data were based on information either recalled during the questionnaire or reconstructed from the household composition at the time of the survey.

The full comparison is shown in Appendix 1. Some important differences were found between the project participants and non-participants. For example, household sizes in project areas were, on average, smaller than those in the comparison areas, and the household heads in project areas were, on average, younger than their counterparts in non-project areas. The average number of adults in project areas was also lower than in non-project areas. On average, land sizes for households in project areas were less than in non-project areas and the number of household members with no education was higher in project households than in non-project households. In addition, the number of households with pit latrines was higher in comparison areas than in intervention areas. There were also significant differences between the households in project areas and non-project areas in terms of whether they grew any crops in 2009 and in their other livelihoods activities in 2009.

These differences that existed before the project have the potential to bias any comparison between the project and comparison groups. It was therefore important to control for these baseline differences when making such comparisons. As described in Section 3, the main approach used in this Effectiveness Review was propensity-score matching (PSM). The full details of the matching procedure applied are described in Appendix 2. After matching, households in the project and comparison groups were

well-balanced in terms of the recalled baseline. In addition, matches could not be found for all of the project participants interviewed: four of the 215 project participant households could not be matched and were dropped from the analysis. The consequence of this is that the estimates of the projects impact presented in Section 5 are not fully representative sample of households in the project communities as they exclude a non-random minority.

All the results described in Section 5 of the report were also tested for robustness by estimating them with various alternative PSM models and linear or probit regression models. The alternative models produced results that are all similar (in size and in statistical significance) to those presented in the tables in this section.

It is important to recall, as highlighted in Section 3, that PSM and regression models can control only for the baseline differences between the households in project and comparison communities for which data was collected in the survey. If there are any 'unobserved' pre-existing differences between the two groups – such as individuals' attitudes, motivation, skills or confidence – then these may bias the estimates of outcomes described in Section 5. The evaluation design and the selection of respondents were intended to minimise any potential for unobserved differences, but this possibility cannot be excluded and must be borne in mind when interpreting the results.

## 5 RESULTS

This report is intended to be free from excessive technical jargon, with more detailed technical information being restricted to the appendices and footnotes. However, there are some statistical concepts that cannot be avoided in discussing the results. In this report, results will usually be stated as the average difference between households living in villages where the project was implemented (that is referred to as the 'intervention group') and the matched households in villages where the project was not implemented (named the 'comparison group'). In the tables of results on the following pages, statistical significance will be indicated with asterisks, with three asterisks (\*\*\*) indicating a p-value of less than 10 per cent, two asterisks (\*\*) indicating a p-value of less than 5 per cent and one asterisk (\*) indicating a p-value of less than 1 per cent. The higher the p-value, the less confident we are that the measured estimate reflects the true impact. Results with a p-value of more than 10 per cent are not considered to be statistically significant.

### 5.1 INTRODUCTION

This section presents a comparison of the project participants and non-participants in terms of various outcome measures relating to the project under review. In the tables of results, asterisks are used to indicate where the differences are statistically significant at least the 10 per cent significance level.

The results are shown after correcting for the baseline and demographic differences found in Section 4.2 using a propensity-score matching procedure. The details of this procedure are described in Appendix 2. All outcomes discussed here have also been tested for robustness using alternative statistical models. In particular, as discussed in Section 4.2, alternative propensity-score matching models were created to take account of whether respondents recalled having been participating in the activities of

the project in 2009. The alternative models produced results that are similar (in size and in statistical significance) to those presented in the tables in this section.

It is important to stress that the results presented in this section are average results across all those who participated in at least one of the interventions that were carried out during the implementation period in seven cooperatives selected under the project up to 2012. Clearly it would be of interest to investigate the effects of the project at a more local level and for specific subgroups – but the small sample sizes limit the potential for detecting any differences between these various subgroups.

One further point that was discussed in Section 3 should be remembered when considering the results presented in this section. The statistical estimation procedures used to derive estimates of outcomes are based only on observable baseline characteristics. If there are any ‘non-observable’ differences between the households surveyed in project and comparison communities – such as individuals’ attitudes or motivation, differences in local leadership, or weather or other contextual conditions – then these may affect the estimates of outcomes. The evaluation design and the selection of respondents were intended to minimise any potential for unobserved differences, but this possibility cannot be excluded and must be borne in mind when interpreting the results.

## 5.2 INVOLVEMENT IN PROJECT ACTIVITIES

Before considering outcome-level changes, it is interesting to consider the proportion of respondents who report exposure to project-related interventions. This is an important consideration as, firstly, it represents the analysis of the first step of the project’s theory of change – i.e. are project participants being exposed to the intended livelihood-support interventions? Secondly, it is important to assess whether respondents in comparison areas also report receiving such support in their communities as this may have an effect on the differences that might be detected between the intervention and comparison groups in the outcome measures reported subsequently. The following activities were considered for the intervention and comparison households from baseline to end line:

- Support in selling coffee to primary coffee cooperative societies (coffee marketing)
- Donation of agricultural inputs (seeds, equipment)
- Donation of jute sacks, slashers, wire mesh
- Support in coffee seedling production, including training

While the project team attempted – to the best of their knowledge – to locate appropriate comparison communities who had not sold their coffee to any primary coffee cooperative society, this was perhaps unlikely to be feasible due to the Ethiopian government policy that requires that coffee farmers be members of cooperative societies. The differences therefore lie in the activities Oxfam implemented with the Limmu Inara Multipurpose Cooperative Union, which other cooperatives selected from the comparison did not implement with their coffee farmers to the best knowledge of the project staff.

**Table 5.1: Involvement in project activities**

	1	2	3	4
	<b>Respondent received support in selling coffee to a cooperative union</b> %	<b>Respondent received agricultural inputs (coffee seed),</b> %	<b>Respondent received support in coffee seedling production</b> %	<b>Respondent received jute sacks</b> %
<i>Intervention group mean:</i>	21.3	6.2	7.1	4.3
<i>Comparison group mean:</i>	1.9	2.2	1.4	1.1
<i>Difference:</i>	19.4*** (3.0)	4.0*** (1.9)	5.7*** (2.0)	3.2** (1.7)
<i>Observations (intervention group):</i>	210	210	210	210
<i>Observations (total):</i>	432	432	432	432

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions and clustered by community

It is important to note that an attempt was made to check whether participation in the activities implemented under the project between the intervention and comparison groups was significantly different. Table 5.1 shows a comparison between those identified as participants (the intervention group) and those identified as non-participants (the comparison group) in terms of their participation in project activities. It should be noted that these figures make a comparison between outcomes after correcting (as far as possible) for the baseline and demographic differences between the intervention and comparison respondents, using the propensity-score matching process described in Appendix 2. Asterisks are used to indicate where the differences are statistically significant at least at the 10 per cent significance level. As is immediately apparent from the table, households that participated in the project are significantly different from the non-participants in terms of the main activities implemented from baseline to the end of the project.

Means and differences shown in the table are those calculated from the propensity-score matching (PSM) kernel model. Asterisks are used to indicate where the differences are statistically significant at least at the 10 per cent significance level.<sup>2</sup>

## 5.3 QUANTITY OF COFFEE PRODUCED AND REVENUES FROM SALES

A key question that this Effectiveness Review sought to answer was whether there was a change in the quantity of coffee produced and sold and the revenues obtained by the farmers who participated in this project. The respondents were asked whether any member of the household had grown coffee in the last 12 months. If yes, they were asked the total quantity of coffee that the household harvested/collected and how much the household sold in the same period and the total revenue obtained from the sales.

Table 5.2 investigated these aspects in the intervention and comparison households selected using kernel propensity-score matching.

Results suggest that, on average, households in the intervention communities produced significantly 64.7 per cent ( $e^{(0.499)-1}$ ) and sold about 1.28 or 128% ( $e^{(0.825)-1}$ ) quintals more coffee than the comparison households. They also received, on average, 1.16 or 116 % ( $e^{(0.773)-1}$ ) higher revenue from the sale of coffee than their counterparts in the comparison communities. These results are statistically significant and also robust to all the estimation models reported in Appendix 3.

**Table 5.2: Quantity of coffee produced and sold**

	1	2	3	4	5	6
	Quantity of coffee produced in the last 12 months (Quintals)	Quantity of coffee produced in the last 12 months (Logarithm of quintals)	Quantity of coffee sold in the last 12 months (Quintals)	Quantity of coffee sold in the past 12 months (Logarithm of quintals)	Revenues from coffee sales in the last 12 months (BIR)	Revenues from coffee sales in the last 12 months (Logarithm of BIR)
<i>Intervention group mean:</i>	1,090.352	6.369	1,727.262	6.388	8,278.619	8.842
<i>Comparison group mean:</i>	506.625	5.870	283.758	5.563	3,427.651	8.069
<i>Difference:</i>	583.727*** (218.847)	0.499*** (0.106)	1,443.504*** (438.740)	0.825*** (0.121)	4,850.968*** (1403.249)	0.773*** (0.159)
<i>Observations (intervention group):</i>	210	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

Table 5.2 therefore provides evidence that the project improved incomes of participatory households through increased quantities of coffee produced and sold coupled with increased revenues from coffee sales. The farmers in the project areas obtained agricultural inputs, including seeds and jute sacks, and were also supported in coffee seedling production. This could have contributed to higher quantities of coffee produced in the intervention areas. They were also able to deliver their coffee to the primary coffee cooperatives as indicated in Table 5.1, which in turn sold the coffee to the main cooperative union, thus increasing coffee sales and revenues in the project areas. This was achieved by ensuring that the primary coffee cooperative societies, after buying coffee from the farmers, sold the coffee to the Limmu Innara Multipurpose Cooperative Union.

## 5.4 PERCEPTIONS OF CHANGES IN: COFFEE PRODUCTION, INCOME, PRICES AND PROCESSING MACHINES

It was also important to assess the farmers' perceptions about changes in total household production of coffee, income from coffee sales, coffee prices and the number of coffee processing machines since 2009. The respondents were asked to state whether their household's total production of coffee had changed since 2009 and whether income from coffee sales had changed. In addition, they were asked whether they had experienced changes in coffee prices and number of coffee processing machines since 2009.

Table 5.3 investigated the impact in the intervention and comparison households using kernel propensity score matching.

**Table 5.3: Changes in coffee production, prices, income and processing machines**

	1	2	3	4
	<b>Increase in total household production of coffee since 2009</b> %	<b>Increase in household income from coffee sales since 2009</b> %	<b>Increase in coffee prices since 2009</b> %	<b>Increase in number of coffee processing machines since 2009</b> %
<i>Intervention group mean:</i>	91.9	81.9	88.1	65.2
<i>Comparison group mean:</i>	82.6	64.6	77.9	41.9
<i>Difference:</i>	9.3*** (3.3)	17.3*** (4.2)	10.2*** (3.6)	23.3*** (4.1)
<i>Observations (intervention group):</i>	210	210	210	210
<i>Observations (total):</i>	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

The estimates in Table 5.3 suggest that 91.9 per cent of the respondents in the intervention group reported an increase in household production of coffee since 2009, compared with 82.6 per cent of the respondents in the comparison group, with a difference of 9.3 percentage points, which is statistically significant different from zero. It appears that 81.9 per cent of the respondents in the intervention group reported an increase in income from coffee sales since 2009, compared with 64 per cent in the comparison group with a significant difference of 17.3 percentage points. The prices of coffee increased by an average of 10.2 percentage points in intervention areas and there is evidence to show that the number of households reporting an increase in the number of coffee processing machines has increased by 23.3 percentage points in intervention areas compared with the non-participant households since 2009. This corroborates evidence regarding quantity of coffee produced and increased revenues

from coffee sales reported in Table 5.2. One important point to note in the table is that a higher proportion of households in the intervention group reported an increase in total household production of coffee as well as increased household income from coffee sales than the comparison group.

As an implementation strategy, Oxfam identified, an organisation called Farm Organic International whose role was to promote coffee marketing on behalf of the Limmu Inara Cooperative Union. The organisation identified export markets for coffee and Oxfam paid for the services of this organisation. These activities could have led to better coffee prices and sales observed in the project areas.

## 5.5 INCREASED USE OF MODERN AGRICULTURAL PRACTICES

Capacity building of coffee farmers was an important activity of this project. This was aimed at ensuring that farmers use modern agricultural practices in coffee production so as to ensure quantity and quality. The use of modern agricultural practices was measured by asking respondents whether they have been using any of the following since the project started: improved seed/seedlings; organic fertilisers; compost manure; sun dryers; cold storage room/collection centre; pest control and caring for trees (column 1).

The outcomes were further disaggregated according to the type of practice. The use of compost manure and organic fertilisers was measured by asking respondents to state whether they practised the use of organic fertilisers and compost manure (column 2). The respondents were also asked to state whether they used either improved seed or seedlings (column 3).

The outcome referring to use of sun dryers or cold storage room was measured by asking respondents to state whether they used any of the three practices (use of greenhouses, sun dryers, and cold storage). In addition, respondents were asked if they applied pesticides (column 4). Lastly, they were asked to state whether they practised drip irrigation or any other type of irrigation (column 5). It is important to note that some of these practices, such as pesticides, were not for coffee husbandry, but for other crops that were also grown by the coffee farmers.

**Table 5.4: Use of modern agricultural practices**

	1	2	3	4	5
	<b>Use of modern agricultural practices (Any of the practices)</b> %	<b>Practised use of compost manure</b> %	<b>Practised use of improved seed/ seedlings</b> %	<b>Practised use of green house, sun dryers and cold storage room for other crops (not for coffee)</b> %	<b>Practised pest control (for other crops and not for coffee production)</b> %
<i>Intervention group mean:</i>	98.6	95.7	91	55.2	77.1
<i>Comparison group mean:</i>	95.2	90.5	87.1	54.1	71.4
<i>Difference:</i>	3.4** (1.6)	5.2** (2.4)	3.9(3.0)	1.1 (5.0)	5.7 (4.3)
<i>Observations (intervention group):</i>	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

Column 1 of Table 5.4 shows the use of modern agricultural practices among the intervention and comparison households. It is important to note that on average there was a 3.4 percentage point increase in the number of participant households reporting use of any of the modern agricultural practices compared with non-participants.

Similarly, on average there was a 5.2 percentage point increase in the number of participant households reporting use of compost manure compared with non-participants (column 2). There was no significant statistical difference observed in the other practices between the intervention and comparison households. It is important to note that these practices were meant for other crops and not for coffee production

One key feature of this finding is that a considerable proportion of coffee farmers in the project areas and non-project areas reported using modern agricultural practices. What may be surprising is why majority of farmers in comparison areas use modern agricultural practices and yet this does not result in higher quantity of coffee produced. It should be remembered that the Limmu Innara Cooperative Union was working with Jimma University in the capacity building of coffee farmers affiliated to the primary coffee cooperative societies. The union organised training for the coffee farmers who were members of primary coffee cooperative societies, and with primary coffee cooperative leaders.

This kind of training was not available to the coffee farmers in the comparison group, and it is therefore likely that coffee farmers in the non-project areas did not get quality training in coffee husbandry that could lead to higher production.

It is also important to note that participant households are in the organic coffee growing area and use of chemicals is not practised. This means they can charge higher prices for their coffee as it is organic.

## 5.6 AGRICULTURAL PRODUCTION

The provision of agricultural inputs, capacity building through training – especially Functional Adult Literacy (FAL) – and the adoption of modern farming methods were just some of the activities performed to increase agricultural production in general and coffee production in particular.

The review used a questionnaire to ask respondents to indicate whether or not they produced a particular agricultural product from a list of 21 different crops in the last 12 months. If they responded positively, they were then asked to estimate the total quantity of each crop produced in the same period. The results for those crops grown by the majority of farmers are presented in Table 5.5

Column 1 indicates the mean quantity of all the crops put together (coffee, maize, mango, banana and chate) between the intervention and comparison households. It is important to note that the quantity of coffee produced had been presented in earlier tables. Results suggest that, on average, households in intervention communities produced 33.7 per cent ( $e^{(0.291)-1}$ ) more of agricultural products in the past 12 months than non-participants. This result is statistically significant and also robust to all the estimation models reported in Appendix A3.

Table 5.5 also compares production between intervention and comparison households on a set of the different agricultural products (columns 2, 3, 4, 5) being among the most widespread products in the sample: mango, maize, banana, and chate. Results suggest that even though there was no significant difference in the quantities of the different agricultural products produced, the project participants produced more than non-participants, except for bananas where the comparison households produced more than the intervention households. The significant increase in total production is most likely to be attributed to significant increase in coffee production shown in Table 5.1 above. This result is also consistent with other estimates reported as robustness checks in Appendix 2. Table 5.5 therefore suggests that there is evidence that the project improved agricultural production in terms of the *total* quantity of agricultural crops produced. The increased use of modern agricultural practices shown in Table 5.3 in the project areas could explain the higher quantity of crops produced in the intervention households.

**Table 5.5: Quantity of agricultural production**

	1		2	3	4	5
	<b>Total quantity of crops produced in the last 12 months (coffee, mango, maize, banana &amp; chate) (Quintals)</b>	<b>Total quantity of crops produced in the last 12 months (Logarithm of quintals)</b>	<b>Quantity of mango produced in the last 12 months (Quintals)</b>	<b>Quantity of maize produced in the last 12 months (Quintals)</b>	<b>Quantity of banana produced in the last 12 months (Quintals)</b>	<b>Quantity of chate produced in the last 12 months (Quintals)</b>
<i>Intervention group mean:</i>	2963.200	7.408	314.167	1095.676	28.233	76.833
<i>Comparison group mean:</i>	1994.411	7.117	48.670	988.738	47.220	27.604
<i>Difference:</i>	968.789** (479.713)	0.291*** (0.104)	265.497 (282.766)	106.938 (226.583)	-18.987* (11.466)	49.229 (48.051)
<i>Observations (intervention group):</i>	210	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

## 5.7 QUANTITY OF SALES FROM AGRICULTURAL PRODUCTION

It was also necessary to investigate whether the project has had any effect on the sales from agricultural crops. For each crop produced, respondents were asked to estimate the quantity sold and the total quantity of crops sold in the last 12 months. Table 5.6 compares the total quantity of crops (coffee, mango, maize, banana and chate) sold in the past 12 months and the quantity of each crop sold. The quantity of coffee sold was shown in Table 5.2.

Table 5.6 provides evidence to show that households in intervention communities sold 94.6 per cent ( $e^{(0.666)-1}$ ) more combined agricultural produce than households in comparison communities. It is important to note, however, that the quantity of each crop sold by intervention communities is not statistically significantly different from the comparison households, even in the case of mangoes where comparison household sold more mangoes than the intervention households. The exception lies with the case of bananas, where non-participant households sold more than the participant households, and results in a statistically significant difference. This could provide further evidence to support the fact that the significant increase in total quantity of crops sold observed between the intervention and comparison household could be due to the significant increase in quantity of coffee sold indicated earlier in Table 5.2. These

findings are in agreement with other estimation models reported in Appendix 3. It is important to note that with coffee marketing enhanced through the membership of the primary coffee cooperative societies, coffee sales in the intervention areas increased compared with the comparison communities who were not selling their coffee to the cooperatives.

It should be noted that there is no statistically significant difference between the intervention and comparison households in the quantity of mangoes, maize and chate produced and sold. This could be explained by the fact that the project focused on improvement of coffee production and sales and so even the training conducted paid particular attention to coffee production and marketing.

**Table 5.6: Quantity of agricultural products sold**

	1	2	3	4	5	6
	<b>Total quantity of crops sold in the last 12 months (coffee, mango, maize, banana and chate) (Quintals)</b>	<b>Total quantity of crops sold in the last 12 months (Logarithm of quintals)</b>	<b>Quantity of mango sold in the last 12 months (Quintals)</b>	<b>Quantity of maize sold in the last 12 months (Quintals)</b>	<b>Quantity of banana sold in the last 12 months (Quintals)</b>	<b>Quantity of chate sold in the last 12 months (Quintals)</b>
<i>Intervention group mean:</i>	2570.248	6.617	8.881	444.786	8.586	175.190
<i>Comparison group mean:</i>	785.477	5.951	34.123	174.812	25.370	11.716
<i>Difference:</i>	1784.771*** (637.295)	0.666*** (0.134)	-25.242 (15.585)	269.974 (265.340)	-16.784** (7.469)	163.474 (105.311)
<i>Observations (intervention group):</i>	210	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

## 5.8 REVENUE FROM AGRICULTURAL SALES

The Effectiveness Review sought to investigate whether the project has had any effect on the revenues arising from agricultural sales, other than coffee. In situations where respondents reported selling any quantity of produce, they were asked to estimate the revenue obtained from selling the quantity of that produce. Table 5.7 compares the total revenue obtained from sales and the revenue obtained from each crop sold in the past 12 months. It is important to note that the total revenue was approximated using all the crops considered here (coffee, mango, maize, banana and chate). Revenues obtained from coffee sales are shown earlier in Table 5.2 and so have not been included here.

The table provides evidence to show that households in intervention communities obtained 92.8 per cent ( $e^{(0.657)-1}$ ) more total revenues from crop sales than households in comparison communities. Since there is no statistically significant difference in the revenues obtained from selling each of the crops, it is most likely that the significant difference in total revenue is due to revenues obtained from coffee sales reported earlier in Table 5.2.

**Table 5.7: Revenue from agricultural sales**

	1	2	3	4	5	6
	<b>Total revenue from sales of crops in the last 12 months (coffee, mango, maize, banana and chate) (BIR)</b>	<b>Total revenue from sales of crops in the last 12 months (coffee, mango, banana and chate) (Logarithms)</b>	<b>Revenue from sale of mango in the last 12 months (BIR)</b>	<b>Revenue from sale of maize in the last 12 months (BIR)</b>	<b>Revenue from sale of banana in the last 12 months (BIR)</b>	<b>Revenue from the sale of chate in the last 12 months (BIR)</b>
<i>Intervention group mean:</i>	8435.324	8.588	39.071	457.048	27.476	141.190
<i>Comparison group mean:</i>	4176.905	7.931	44.966	779.422	74.421	160.446
<i>Difference:</i>	4258.419*** (1415.059)	0.657*** (0.169)	-5.895 (20.735)	-322.374 (224.890)	-46.945 (37.472)	-19.256 (132.026)
<i>Observations (intervention group):</i>	210	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

## 5.9 OVERALL HOUSEHOLD INCOME

Measuring household income directly is problematic: self-reported measures of total income are generally regarded as unreliable, given the wide variety of endeavours such populations engage in to generate income. Most households were engaged in other livelihood activities; a direct income measure would have to collect detailed information about the contribution of each of these activities to household income.

For these reasons, the survey did not attempt to collect data on total household income directly. However, there is a widely recognised and strong association between household income and consumption.<sup>3</sup> The Effectiveness Review therefore followed common practice in micro-level socio-economic analysis, by considering household consumption and expenditure as an indicator of income.

To that end, respondents were asked to provide detailed information about their recent expenditure on both food and non-food items. Firstly, the respondents were asked from a list of 26 products what types of food they consumed over the previous seven-day

period, as well as the particular quantities. The quantities of each food item consumed were then converted into a monetary value. This was done by asking the respondent how much was paid for the food item in question or – if the food item was from the household’s own production – how much it would be worth if it was purchased from the local market. The respondents were also asked how much they spent on particular regular non-food items and services from a list of 19 items, such as fuel, toothpaste, and transport fares over the past four weeks. Finally, they were asked to estimate the value of other occasional types of expenditure that they incurred over the previous 12 months from a list of 25 items, which included clothes, medical expenses and home repair. The household expenditure measure was calculated by converting each of the expenditure types into a per-day per-capita<sup>4</sup> figure and adding them together. This figure was then divided by a factor representing household size, to generate a per-day, per-person expenditure figure. As with the measures of agricultural sales, the expenditure variable has been expressed on a logarithmic scale, to reduce the influence on the overall result of any households with extreme values for total consumption. The comparison of expenditure between supported households and comparison households, both before and after logarithmic transformation, is shown in Table 5.8.

It can be seen in column 1 of the table that the value of food consumed within the households of project participants and non-participants was not significantly different. However, the logarithmic transformations did show a significant difference (column 2).

**Table 5.8: Household consumption**

	1	2	3	4	5
	<b>Food consumption per adult equivalent per day (Pesos)</b>	<b>Food consumption per adult equivalent per day (Logarithm of pesos)</b>	<b>Total household consumption per adult equivalent per day (Pesos)</b>	<b>Total household consumption per adult equivalent per day (Logarithm of pesos) - New Global Indicator</b>	<b>Old Global Indicator (Logarithm - Global indicator for livelihoods support)</b>
<i>Intervention group mean:</i>	11.493	2.251	23.928	2.974	0.710
<i>Comparison group mean:</i>	10.554	2.123	18.501	2.727	0.498
<i>Difference:</i>	0.939 (0.765)	0.128** (0.060)	5.427*** (1.491)	0.247*** (0.059)	0.212*** (0.049)
<i>Observations (intervention group):</i>	210	210	210	210	210
<i>Observation (total):</i>	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

The analysis finds evidence to suggest that intervention households report statistically significant different levels of daily per capita consumption (Columns 3 and 4) compared to comparison households.

The food consumption per adult equivalent, per day, was transformed into logarithms and this created a 13.6 per cent ( $e^{(0.128)}-1$ ) difference between the intervention and comparison households. Similarly, total household consumption per adult equivalent per day (overall household income) was transformed into logarithms in order to interpret the percentage change between the intervention and comparison households for this indicator. There was a 28.01 per cent ( $e^{(0.247)}-1$ ) change in the overall household income as measured by consumption and expenditure as illustrated above. This measure also represents Oxfam GB's New 'Global Indicator' for livelihoods Effectiveness Reviews. In past years this indicator was measured as the proportion of the samples whose daily consumption exceeds the median of the unmatched comparison sample. In order to provide some continuity with previous Effectiveness Reviews column 5 in Table 5.8 also provides estimates for the old measure. This shows that 71 per cent of households in the intervention sample had expenditures exceeding the critical value, compared to 49.8 per cent of the matched comparison sample, giving a difference, attributable to the project, of 21.1 percentage points. This is statistically significant at the 1 per cent level. All these estimates were statistically significant different from zero and robust to other estimation models in Appendix 3.

In summary, the results of Table 5.8 imply that household consumption was significantly higher, on average, among the households of project participants than among households of non-participants at the time of the survey.

## 5.10 HOUSEHOLD ASSET WEALTH

An alternative way to consider income is to investigate asset ownership. Thus respondents were asked about their ownership of various types of household goods, productive assets and livestock, as well as about the condition of their housing. These data were used to create a wealth index using Cronbach's alpha.<sup>5</sup> A total of 36 assets and other wealth indicators were used to construct the household wealth index with their inter-item correlations. The wealth indices were then created applying principal component analysis (PCA) to the selected indicators. PCA is a data reduction technique that narrows in on the variation in household asset ownership that is assumed to represent wealth status: the more an asset type is correlated with this variation, the more weight it is given.

To examine changes in wealth indicators, the data about ownership of particular assets and housing characteristics were aggregated into a single index of material wealth. To do this, they were first checked to see which of the indicators appeared to be good indicators of wealth. If every asset and housing characteristic is an indicator of household wealth, they should be all correlated with each other. That is, a household that scores favourably on one particular wealth indicator should be more likely to do so for other wealth indicators. Items that had low correlations or negative correlations with the others were therefore not included in the index. In particular, ownership of most types of livestock was not found to be correlated with ownership of other assets, nor was ownership of a motorbike or car, nor was some other minor characteristics. These variables were therefore not used in the construction of the wealth index.

Two separate wealth indices were created. The wealth index based on the data recalled from 2009 and at the time of the survey. The 2009 measure is shown in the table of summary statistics in Appendix 1, and has been used throughout this analysis to control (to the greatest extent possible) for baseline differences in wealth status among the households of the various treatment groups. The second index allows the wealth indicators at the time of the survey to be compared between the intervention and comparison households (Table 5.9, column 1).

A comparison of the participant and non-participant households in terms of the wealth index at the date of the survey is shown in column 1 of Table 5.9. Positive numbers represent a positive change in wealth indicators. It can be seen in the table that the difference in wealth index between the two groups was statistically significantly different from zero.

**Table 5.9: Wealth index**

	1	2
	Wealth index at date of survey	Change in wealth index between 2009 and date of survey (scale from -1 to 1)
<i>Intervention group mean:</i>	0.087	0.019
<i>Comparison group mean:</i>	0.052	-0.167
<i>Difference:</i>	0.035*** (0.012)	0.186 (0.124)
<i>Observations (intervention group):</i>	210	210
<i>Observations (total):</i>	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

Column 2 shows the figures for the change in wealth index for project and comparison households between 2009 and the date of the survey. The change in wealth index is scaled so that a household that saw no changes in wealth indicators has a score of zero, while the household that saw the greatest change in wealth indicators has a score of 1. (The household that experienced the largest *negative* change in wealth index has a score of approximately -0.9 on this scale.) It can be seen that project participants on average experienced improvements in their wealth indicators since 2009 (that is, the changes in wealth index were positive). The comparison households, on the other hand, experienced no improvements in their wealth indicators since 2009 (that is, the changes in wealth index were negative). However, the difference between the two was not statistically significantly different from zero.

## 5.11 CARE MODULE

### Measurement and assessment of care work in impact evaluation

This section aims to summarise the measurement approaches for project's impact evaluation on care. Tools and indicators were designed to look at differences at individual/household level and were intended to be used in surveys where the respondent is a woman. In livelihoods surveys where the respondent is not necessarily a woman, the suggested approach is to create a section at the end of the questionnaire in which these questions can be addressed by any woman in the household. This is what was applied in this impact evaluation of care.

There were two different components that were collected and taken in to account in the measurement of care. The first one is *time* devoted to care activities. The second one is *intensity* and *effort* devoted to care. Moreover, when performing evaluation of projects, it is also important to define a mechanism through which change may happen. Three possible channels are: redistribution of care responsibilities with other household members; acquisition of time and labour saving equipment; change in opinion and awareness. The impact evaluation followed the three channels in order to try to check whether there were significant differences between the intervention and the comparison groups

### **5.11.1 Time use on care**

There are (at least) two approaches for assessing project's impact on care **time use**. The first one is estimating the number of hours spent during each day in care activities, and comparing the intervention and comparison groups. The second one is assessing self-reported change in time spent in care activities since the programme began, and comparing intervention and comparison groups. In this evaluation both approaches were used.

### **5.11.2 Hours spent in a day on care and productive activities**

When assessing care, estimating precisely the time devoted to particular activities is extremely important. However, this can be difficult and time consuming. The tools to use vary depending on the relative weight care is given in the project's theory of change and therefore the relative time in the questionnaire.

In projects where care activities play an important role, it is desirable to choose impact evaluation analysis and data collection processes that are more accurate, although also more time consuming. This is the approach presented in this report.

The recommended, and more precise, option to estimate the total amount of hours spent doing care is to ask the respondent to report what she was doing in each hour during the last 24 hours. The enumerator then reports the code that corresponds to the mentioned activity. The respondent is also asked if she was responsible for looking after a child or dependent adult at the same time in which the main activity was reported.

At data analysis level, every time a code for care activities is reported, the hour will be added in order to estimate the total number of hours each respondent spent involved in care activities. In addition to this, hours will also be included for the counting if a woman reported doing non-care activities as main activity, but said they were also responsible for looking after a child or dependent adult. Primary care hours refer to the following activities: shopping (including walking to the market); washing, drying, ironing and mending clothes; food and drink preparation; grinding and pounding; cleaning; fuel collection; water collection; child care; dependent adult care; care of the disabled; care of community members.

Primary and secondary care hours refer to hours where there was any care responsibility. Sleep hours refers to the number of hours taken by women to sleep. No-work hours include: doing nothing; sleeping; personal care and eating; attending school or training; travelling for other reasons; leisure time; and religious activity. Multitask care hours refer to doing at least two care activities at the same time

**Table 5.10: Time taken by women for all activities and for care work**

	1	2	3	4	5	6
	<b>Primary care hours</b>	<b>Time taken by women caring for children and elderly in addition to other activities %</b>	<b>Primary and secondary care hours</b>	<b>Hours taken to sleep</b>	<b>No-work hours</b>	<b>Multi-tasking care hours</b>
<i>Intervention group mean:</i>	3.757	16.2	4.310	4.052	6.100	4.500
<i>Comparison group mean:</i>	5.374	26.1	6.061	5.618	8.298	6.640
<i>Difference:</i>	-1.617*** (0.412)	-9.9*** (4.2)	-1.751*** (0.472)	-1.566*** (0.436)	-2.198*** (0.602)	-2.140*** (0.489)
<i>Observations (intervention group):</i>	210	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

The results in column 1 of Table 5.10 provide evidence to show that, on average, the total time taken by women doing primary care activities in a 24-hour period is significantly higher for women in comparison households than women in intervention households. Women in comparison households take 1 hour 36 minutes more to perform primary care activities compared to their counterparts in the project areas.

Column 2 records the time taken by women to care for children and dependent adults in addition to other activities. Respondents were asked whether they were responsible for looking after a child or dependent adult during the hour they were performing other duties. Again the results indicate that, on average, women in comparison households took about 9.9 percentage points more of their time caring for children and elderly in addition to other activities than their counterparts in the intervention areas. These findings are consistent with the ones obtained using other estimation models shown in Appendix 2.

Column 3 compares primary and secondary care hours between the intervention and comparison areas. It is clear that, on average, women in comparison areas take 1 hour and 45 minutes longer to do any care activity than women in project areas and this is statistically significantly different from zero.

Interestingly, women in project areas take less time to sleep, on average (1 hour 33 minutes less), than women in the non-project areas (column 4). Similarly, women in comparison areas take 2 hours 11 minutes more, on average, doing no work than women in project areas (column 5). These two findings are consistent with women in

project areas take less time to sleep and less time doing nothing, on average, than women in comparison areas.

Another significant finding is that women in comparison areas take 2 hours 8 minutes more, on average, to do multitasking care activities than women in intervention areas. This result is significantly different from zero and robust to other estimation models presented in Appendix A3.

### 5.11.3 Redistribution of responsibilities for care within the household

The first mechanism that can lead to changes in time and intensity of the care activities in women’s lives is a more equal **redistribution of responsibilities within the household**. Most of Oxfam’s projects aim to change the power structures and gender norms in the favour of women. A suggested indicator to investigate gender roles and domestic responsibilities for care is to investigate if the amount of time that men and boys in the household spend on care activities had increased or decreased since the beginning of the project.

It is worth remembering that changes over time can be causally attributed to the project intervention if estimates identified in the intervention group differ from estimates from a counterfactual.

In addition to this indicator, the amount of time that other women in the household spend on these activities is investigated to find out if it had increased or decreased since the project began. This information is intended to investigate potential unintended negative effects that might appear when care responsibilities are simply passed over to other female household members without a change in domestic roles.

**Table 5.11: Change in time spent by men, women or other women in household to care for children, elderly or other household members**

	1	2	3
	<b>Time spent by men in the household to care for children, the elderly or other members of household increased</b> %	<b>Time spent by women to care for children, the elderly or other members of household increased</b> %	<b>Time spent by other women in the household to care for children, the elderly or other members of household increased</b> %
<i>Intervention group mean:</i>	35.7	38.1	32.9
<i>Comparison group mean:</i>	22.2	29.3	23.7
<i>Difference:</i>	13.5*** (4.4)	8.8* (4.7)	9.2* (3.7)
<i>Observations (intervention group):</i>	210	210	210
<i>Observations (total):</i>	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

**Table 5.12: Time spent by men, women or other women in household to undertake housework increased**

	1	2	3
	Time spent by men to undertake housework increased %	Time spent by women to undertake housework increased %	Time spent by other women in hh to undertake housework increased %
<i>Intervention group mean:</i>	31.4	36.2	32.9
<i>Comparison group mean:</i>	20.3	34.8	21.9
<i>Difference:</i>	11.1** (3.6)	1.4 (4.0)	11** (3.6)
<i>Observations (intervention group):</i>	210	210	210
<i>Observations (total):</i>	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

**Table 5.13: Time spent by men, women or other women in hh to carry out income generating activities increased**

	1	2	3
	Time spent by men to undertake income generating activities increased %	Time spent by women to undertake income generating activities increased %	Time spent by other women in hh to undertake income generating activities increased %
<i>Intervention group mean:</i>	37.6	37.6	32.4
<i>Comparison group mean:</i>	38.8	38.8	27.0
<i>Difference:</i>	-1.2 (5.1)	-1.2 (5.1)	5.4 (5.0)
<i>Observations (intervention group):</i>	210	210	210
<i>Observations (total):</i>	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

**Table 5.14: Time spent sleeping by men, women or other women in hh increased**

	1	2	3
	Time spent by men to sleep increased	Time spent by women to sleep increased	Time spent by other women in hh to sleep increased
	%	%	%
<i>Intervention group mean:</i>	14.3	15.2	13.8
<i>Comparison group mean:</i>	13.4	17.1	11.3
<i>Difference:</i>	0.9 (2.9)	-1.9 (3.1)	2.5 (2.8)
<i>Observations (intervention group):</i>	210	210	210
<i>Observations (total):</i>	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

The four activities listed were: responsibility for the care of children, elderly or other household members; housework (e.g. cooking, cleaning, washing etc); income generating activities (e.g. farming, business, petty trade, handicrafts, etc.); and sleeping at night.

Women were asked whether the amount of time **they** spent on each of the four activities increased or decreased since the project started. They were also asked whether the amount of time **men** in the household spent on each of the four activities had increased or decreased since the beginning of the project. Finally, women were asked to state whether the amount of time **other women in the household** spent on each of the four activities had increased or decreased since the beginning of the project. The results are presented in Tables 5.11, 5.12, 5.13 and 5.14.

Table 5.11 shows the change in amount of time spent by women, men or other women in the household on care of children, the elderly or other members of the household. Results indicate that since 2009, the amount of time spent by men on care increased, by about 13.5 percentage points, on average, in the intervention households more than in the comparison households. Similarly, the amount of time spent by women on care was reported to have increased by about 8.8 percentage points, on average, in the intervention areas more than the comparison communities. In addition, the amount of time spent on care by other women in the household was reported to have increased by 9.2 percentage points percentage points, on average, in intervention areas more than the comparison areas.

Table 5.12 shows the change in the amount of time spent by women, men or other women in the household on housework. There is evidence to show that the amount of time spent by men on housework in the intervention areas increased by about 11.1 percentage points, on average, compared with the amount of time spent on the same activity by men in the comparison communities. The amount of time spent by other women in the household on housework also increased by 11 percentage points in intervention households more than the amount of time spent on the same activity by women in comparison communities. However, there was no significant difference in the amount of time spent on housework by women in the intervention and comparison households.

Table 5.13 and 5.14 show the change in the amount of time spent by men, women or other women in the household on income-generating activities and sleeping at night respectively. There is no evidence to show that the amount of time spent by men, women or other women in the household was significantly different for the two activities in the intervention and comparison households.

#### 5.11.4 Ownership of time/labour-saving equipment

The second mechanism through which we can experience a reduction in women's time and intensity of work devoted to care activities is through the acquisition of **time/labour-saving equipment**. This indicator is intended to capture ownership and women's decision-making control over these types of equipment. It was captured by asking each respondent if her household owned assets from a given list of items. If the answer was affirmative, it was then asked who had decision-making control over the item – that is, who could decide whether to sell it, if necessary.

The list of equipment included: charcoal stove; energy-saving stove; grinding mill; dustbin; compost pit; water piping; mashing facilities; and rainwater harvesting equipment. The list could be adapted and changed, depending on the local context, in order to reflect relevant time-saving activities in the area under analysis.

**Table 5.15: Ownership of labour saving equipment**

	1	2	3	4	5
	Ownership of labour saving equipment %	Women have decision-making control over the equipment %	Men have decision-making control over the equipment %	Respondent and husband jointly have decision-making control %	Respondent and other household members jointly have decision making control %
<i>Intervention group mean:</i>	43.8	26.2	1.9	14.8	8.6
<i>Comparison group mean:</i>	55.8	42.1	10.1	15.6	7.5
<i>Difference:</i>	-12.0** (5.0)	-15.9*** (4.6)	-8.2*** (2.5)	-0.8 (3.6)	1.1 (2.8)
<i>Observations (intervention group):</i>	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

From the results in Table 5.15, there is evidence to show that 12 percentage points more households in the comparison areas had labour saving equipment, on average, compared with their counterparts in the project areas. In addition, about 15.9 percentage points more women in the comparison communities, on average, have decision-making control over the equipment compared with women in the intervention households. About 8.2 percentage points more men on average have decision-making control over equipment in comparison households compared with intervention households. However, there is no evidence to show that either respondent and husband jointly or respondent and other household members jointly have decision-

making control over the equipment for either the intervention or comparison households.

### 5.11.5 Attitude towards and awareness of care work

The last mechanism through which change in time and intensity for care can take place is through differences in **attitude toward care work**. This indicator is captured by asking the respondents to what extent they agree or disagree with the following statements:

- Women are by nature better at housework and looking after children and dependent adults than men.
- Men are by nature better at providing for their household than women.
- Husbands should help their wives with housework and looking after children and dependent adults.
- Women and men should decide together about who does what in the household.
- The government should help households with housework and looking after children and dependent adults (e.g. providing better access to water or child care).
- Housework and looking after children and dependent adults requires significant skills.

As for other indicators, this list can be reduced to a minimum of two questions, depending on the relevance of the indicator and time availability for the respondent.

**Table 5.16: Subjective statements on care, housework, providing for the family and decision making**

	1	2	3	4	5	6
	Agree that women are by nature better at housework and looking after children and dependent adults than men	Agree that men are by nature better at providing for their household than women	Agree that husbands should help their wives with housework and looking after children and dependent adults	Agree that women and men should decide together about who does what in the household	Agree that government should help households with housework and looking after children and dependent adults	Agree that housework and looking after children and dependent adults require significant skill
<i>Intervention group mean:</i>	46.2	38.6	46.2	44.3	47.1	47.1
<i>Comparison group mean:</i>	63.6	53.2	65.4	62.9	68.2	65.8
<i>Difference:</i>	-17.4*** (5.0)	-14.6*** (0.051)	-19.2*** (4.9)	-18.6*** (4.9)	-21.1*** (4.8)	-18.7*** (5.0)
<i>Observations (intervention group):</i>	210	210	210	210	210	210
<i>Observations (total):</i>	642	642	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

Table 5.16 shows the differences in the attributes. Column 1 shows that about 17.4 percentage points more respondents in comparison households, on average, agree that women are by nature better at house work and looking after children and dependent adults than respondents in the intervention households. In addition, about 14.6 percentage points more respondents in comparison households on average agree that men are by nature better at providing for their households than women (column 2). There is evidence to show that 19.2 percentage points more respondents in non-participant households, on average, agree that husbands should help their wives with house work and looking after children and dependent adults than their counterparts in the participant households (column 3). Similarly, 18.6 percentage points more respondents in the comparison households, on average, agree that women and men should decide together about who does what in the household than respondents in the participant households (column 4).

In the same vein, 21.1 percentage points more respondents in the non-participant households, on average, agree that government should help households with house work and looking after children and dependent adults than respondents in participant households (column 5). Likewise, about 18.7 percentage points more respondents in non-participant households, on average, agree that house work and looking after children and dependent adults require significant skill than respondents in participant households (column 6).

These findings tend to suggest that women in non-participant households are overburdened with housework and care work compared with their counterparts in the participant households. It could be that WeCare programme interventions in participant households may be bearing fruit, but this needs further investigation in qualitative studies.

## 6 CONCLUSIONS

### 6.1 CONCLUSIONS

This Effectiveness Review has found evidence that the coffee value chain project has made a positive contribution to the livelihoods of rural coffee producers. There were significant differences between project participants and non-participants in terms of quantity of coffee produced and sold, and the corresponding revenues obtained.

This finding was reinforced by project participants themselves, who also indicated that coffee prices and household income from coffee sales have increased since 2009. On average, the quantity of coffee produced increased by 64.7 per cent among participant households more than non-participants. The quantity of coffee sold on average increased by 1.28 times or 128 per cent in households that participated in the project more than non-participants.

Another question that the Effectiveness Review sought to answer was whether there was a change in revenue arising from the quantities of coffee sold. Survey data showed that on average, revenues increased by about 1.16 times or 116 per cent among project participants compared with non-participants.

In addition, on average, the total quantity of crops harvested and sold among participant households was 33.7 per cent and 94.6 per cent higher, respectively, than non-participants. The revenue arising from sales of total quantity of products was 92.8 per cent higher on average among project participants than non-participants.

Another key conclusion from this Effectiveness Review is that on average, there was a 3.4 percentage point greater increase in the number of participant households reporting use of modern agricultural practices than non-participants. In addition, there was a significant difference of about 5.2 percentage points between participant households and non-participants regarding use of compost manure. It should be noted that these practices ensured quantity and quality of production among project participants

One important conclusion from this Effectiveness Review is that on average, consumption is 13.6 per cent higher among project participants. The difference in household consumption between the project participants and those of the matched non-participants, combined with the increase in wealth indicators (asset ownership, livestock and housing conditions) since 2009, suggests that net household income among the participants is significantly higher than it would have been without this project. The estimates derived from the measures of household consumption and expenditure implies that the magnitude of this increase in income is about 28 per cent and is statistically significantly different from zero. This measure also represents **Oxfam GB's New 'Global Indicator' for livelihoods** Effectiveness Reviews.

In the past years this indicator was measured as the proportion of the samples whose daily consumption exceeds the median of the unmatched comparison sample. In order to provide some continuity with previous Effectiveness Reviews estimates for the old Global Indicator was also provided. .

The old Oxfam GB Global Indicator for livelihoods support, only captures increases in income that cross the threshold (the median income among the comparison group). A very poor household (in the lower half of the income distribution) may have increased their income substantially, but this is not relevant for the global indicator unless they reach the threshold. Similarly, a household that would have been in the top half of the income distribution (i.e. with above-average income) would never be counted towards the indicator. Only households that started below the threshold and experienced an increase in income to take them above the threshold (median in the comparison) are currently counted positively.

In this Effectiveness Review, there was a 21.2 percentage points change in average household income (as measured by household consumption and expenditure per adult equivalent person, per day) for the **old Global Indicator** for Livelihoods.

## 6.2 PROGRAMME LEARNING CONSIDERATIONS

### **Strengthening marketing channels in livelihood programmes leads to better incomes for farmers.**

The results of this Effectiveness Review add to the proof that farmers' markets provided by the coffee cooperative unions have a positive impact in terms of increasing income. This Effectiveness Review provides clear evidence that the new marketing channels for coffee farmers (selling through the cooperatives), which cut out intermediaries, and the related support provided to producers under this project, have led to a significant improvement in household income. This evidence can be used to strengthen the case both for scaling up the farmers' markets in Ethiopia and for encouraging adoption in other parts of Ethiopia where coffee is grown.

### **Engaging stakeholders with clearly defined roles should be considered.**

The critical role of the stakeholders in moving the project success forward, through engaging them from the outset, contributed to positive outcomes. One key learning point from this Effectiveness Review is the pivotal role played by the bureaux and sectoral institutions, such as Jimma University and the cooperatives agency in delivering an integrated support for success, by influencing, and by demonstration and scaling. Jimma University led in capacity-building training to equip coffee farmers with the necessary skills for coffee production. The cooperatives unions facilitated the marketing of coffee and ensured that the farmers sold greater quantities and obtained higher revenues for their produce.

### **Capacity building of beneficiaries both before and during project implementation should be encouraged in livelihoods interventions.**

The Functional Adult Literacy (FAL) scheme, spear-headed by Limmu Inara Cooperative Union, was aimed at building the capacity of coffee farmers particularly in the use of modern agricultural practices, with a view to improving the quantity and quality of coffee produced. Indeed, there is evidence from this Effectiveness Review that farmers adopted modern methods of farming. The Integrated Functional Adult Learning contributed to the acquisition of new skills and knowledge and should therefore be encouraged in future implementation strategies in projects of a similar nature. Leveraging on external stakeholders and global events for demonstration, synergy and influencing, such as the global PS learning events organised by Oxfam, Catholic Relief Services (CRS) and Slow Food and a finance fare at BioFach, should also be encouraged in implementing future projects.

### **The identification of stages within a given value-chain where women can actively engage should be considered.**

One important learning consideration is that when designing a project, it is important to consider in which stages of the value chain women best fit and what kind of special support they may require. At times, it may not matter whether the commodity is a cash crop mainly dominated by men. However, knowing which area works best for women is crucial. In this particular coffee value chain project, women engaged in coffee seedling preparation. This proved to be quite relevant for them and consequently more seedlings were produced, which ultimately boosted coffee production.

### **Institutional strengthening of the Limmu Inara Cooperative Union worked better than support through a business service provider.**

Initially, Oxfam commissioned a business service provider to support the Limmu Inara Cooperative Union to play a facilitation role and link them to the export buyers. One of the lessons learnt is that strengthening the institutional capacity of the union can make it more effective in discharging its functions. The union hired its own commercial manager to facilitate its own export link. As a result, the union managed to directly connect to export buyers and was able to negotiate for prices. This ensured that the union had more potential buyers who could offer better coffee prices. Consequently, coffee farmers were able to sell more of their produce at better prices and hence obtain higher revenues.

# APPENDIX 1: BASELINE STATISTICS BEFORE MATCHING

**Table A1.1: Descriptive statistics: comparison between intervention and comparison households at baseline**

	Descriptive Statistics				
	N	Intervention mean	Comparison mean	Difference	t-statistic
Household size	647	5.060	6.074	-1.014***	0.189
Age of household head	647	41.688	49.023	-7.336***	1.233
Main work of HHH – farming	647	1.688	1.098	0.590***	0.153
Income source of household – farming	647	0.840	0.981	-0.141***	0.026
Income source – household – livestock	647	0.097	0.112	-0.014	0.025
Income source – household – business	647	0.039	0.060	-0.021	0.018
Number of adults in household	647	2.731	3.260	-0.529***	0.104
Household members – no education	647	0.442	0.316	0.126***	0.041
Households toilet is pit latrine	647	0.690	0.833	-0.143***	0.036
Household member own dwelling	647	0.887	0.981	-0.095***	0.023
wealth_index_2009_q5_2	647	0.220	0.158	0.062*	0.033
wealth_index_2009_q5_3	647	0.192	0.219	-0.026	0.033
wealth_index_2009_q5_4	647	0.157	0.284	-0.126***	0.033
wealth_index_2009_q5_5	647	0.174	0.251	-0.078**	0.033
Number of cows owned by household	647	1.389	1.456	-0.067	0.203
Number of oxen owned by household	647	0.880	1.028	-0.148	0.123
Number of sheep owned by household	647	0.377	0.530	-0.153	0.117
Any member of household grew crops	647	0.808	0.907	-0.099***	0.030
Any household land irrigated	647	2.231	2.465	-0.234***	0.088
Household had agricultural land	647	0.871	1.054	-0.183	0.135
Land area farmed	647	1.237	1.700	-0.463**	0.216
Time to reach centre of community	647	31.634	21.619	10.016***	3.106
<b>N</b>	<b>647</b>				

t statistics in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01;

## APPENDIX 2: METHODOLOGY USED FOR PROPENSITY-SCORE MATCHING

The analysis of outcome variables presented in Section 5 of this report, involved group mean comparisons using propensity-score matching (PSM). The basic principle of PSM is to match each participant with a non-participant that was observationally similar at baseline and to obtain the programme treatment effect by averaging the differences in outcomes across the two groups after project completion. Unsurprisingly, there are different approaches to matching, i.e. to determining whether or not a household is observationally 'similar' to another household. For an overview, we refer to Caliendo and Kopeinig (2008).<sup>6</sup>

The following sections describe and test the specific matching procedure followed in this Effectiveness Review.

### Estimating propensity scores

Given that it is extremely hard to find two individuals with exactly the same characteristics, Rosenbaum and Rubin (1983) demonstrate that it is possible to match individuals using a prior probability for an individual to be in the intervention group, naming this its *propensity score*. More specifically, propensity scores are obtained by pooling the units from both the intervention and comparison groups and using a statistical probability model (e.g. a *probit* regression) to estimate the probability of participating in the project, conditional on a set of observed characteristics.

Table A2.1 presents the probit regression results used to estimate the propensity scores in our context. To guarantee that none of the matching variables were affected by the intervention, we only considered variables that were measured at baseline, and only those variables that were unlikely to have been influenced by anticipation of project participation (Caliendo, 2008).

**Table A2.1: Estimating the propensity score**

Covariates	Coefficients	Standard Error	P>z
Household size	.0281418	.0102249	0.006
Age of household head	.0077537	.0014926	0.000
Main work of HHH – 2009	-.0458187	.018202	0.013
Income source of household – farming	.2067238	.0607671	0.015
Income source – household – livestock	-.0512406	.0583592	0.401
Income source – household – business	.1230355	.1046811	0.214
Number of adults in household	-.0059083	.0191025	0.757
Household members – no education	-.1741492	.0390825	0.000
Households toilet is pit latrine	.0451855	.049308	0.370
Household member own dwelling	.1855889	.062405	0.029
wealth_index_2009_q5_2	.0591034	.0736916	0.411

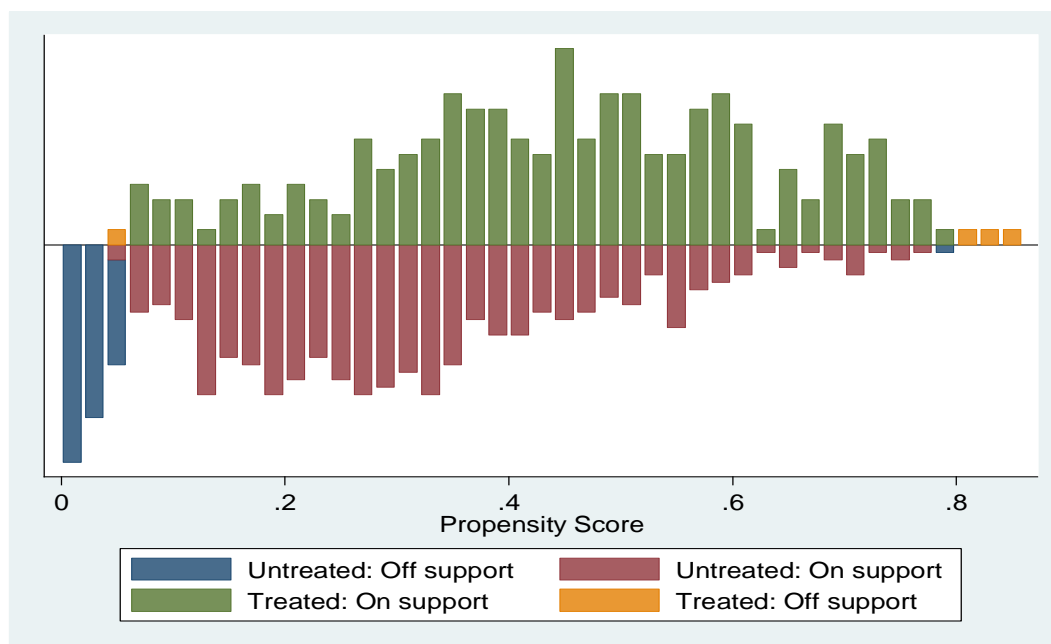
wealth_index_2009_q5_3	.1619507	.0796228	0.034
wealth_index_2009_q5_4	.216421	.0841902	0.007
wealth_index_2009_q5_5	.1370157	.0928935	0.124
Number of cows owned by household	-.0206683	.0096914	0.033
Number of oxen owned by household	-.0114178	.0139568	0.413
Number of sheep owned by household	.0050045	.0138093	0.717
Any member of household grew crops	-.1543686	.0840751	0.053
Any household land irrigated	.0318419	.021574	0.139
Household had agricultural land	.0064501	.0115209	0.642
Land area farmed	.0044539	.0087952	0.612
Time to reach centre of community	-.0019365	.0008084	0.018
<b>Observations</b>	<b>647</b>		

## Defining the region of common support

After estimating the propensity scores, the presence of a good *common support area* needs to be checked. The area of common support is the region where the propensity-score distributions of the treatment and comparison groups overlap. The common support assumption ensures that ‘treatment observations have a comparison observation “nearby” in the propensity score distribution’ (Heckman, LaLonde and Smith, 1999). Since some significant differences were found between the intervention and comparison groups in terms of the baseline and demographic characteristics (as detailed in Section 4.2), some of the households in the intervention group are too different from the comparison group to allow for meaningful comparison. We developed a minima and maxima comparison, deleting all observations whose propensity score was smaller than the minimum and larger than the maximum in the opposite group (Caliendo, 2008). In this particular case, none of the 432 households surveyed in the comparison villages and 5 of the 215 households surveyed in the intervention villages were dropped because of lying outside the common support area. This means that the estimates of differences in outcome characteristics between the various treatment groups only apply to those intervention households that were not dropped; that is, they do not represent the surveyed population as a whole.

Figure A2.1 illustrates the area of common support and indicates the proportion of households lying on and off the common support area, by treatment group.

Figure A2.1: Propensity score on and off common support



## Matching intervention households to comparison households

Following Rosenbaum and Rubin (1983), after estimating the propensity scores and defining the area of common support, individuals are matched on the basis of their propensity score. The literature has developed a variety of matching procedures. For the main results presented in this Effectiveness Review we chose to employ the method of kernel matching (note that we use alternative matching procedures as a means of robustness checks in Appendix 3). The kernel matching method weights the contribution of each comparison group member, attaching greater weight to those comparison observations that provide a better match with the treatment observations. One common approach is to use the normal distribution with mean zero as a kernel, and weights given by the distribution of the differences in propensity score. Thus ‘good’ matches get a larger weight than ‘poor’ matches.

We used the *psmatch2* module in STATA using 0.06 as a bandwidth and restricted the analysis on the area of common support. When using PSM, standard errors of the estimates were bootstrapped (stratified by community) using 1,000 repetitions to account for the additional variation caused by the estimation of the propensity scores and the determination of the common support.<sup>7</sup>

## Check balancing

For PSM to be valid, the intervention group and the matched comparison group need to be balanced in that they needed to be similar in terms of their observed baseline characteristics. This should be checked. The most straightforward method to do this is to test whether there are any statistically significant differences in baseline covariates between the intervention and comparison group in the matched sample. The balance of each of the matching variables after kernel matching is shown in Table A2.2. None of the variables implemented for the matching is statistically significant once the matched sample is used.

**Table A2.2: Balancing test on the full set of baseline covariates**

Psmatch2 Treatment Assignment	Psmatch2 Common Support		Total
	Off Support	On Support	
Untreated	0	432	432
Treated	5	210	220
<b>Total</b>	<b>5</b>	<b>642</b>	<b>652</b>

Covariates	Unmatched Matched	Mean		% bias	% reduction bias	P-value
		Treated	Control			
Household size	<b>U</b>	6.0744	5.0602	44.1		0.000
	<b>M</b>	6.0238	5.9655	2.5	94.3	0.791
Age of household head	<b>U</b>	49.023	41.688	50.3		0.000
	<b>M</b>	48.686	48.99	-2.1	95.9	0.831
Main work of HHH-2009	<b>U</b>	1.0977	1.6875	-35.9		0.000
	<b>M</b>	1.1	1.1145	-0.9	97.5	0.861
Income source of household-farming	<b>U</b>	0.9814	0.84028	51.0		0.000
	<b>M</b>	0.9805	0.9777	1.2	97.7	0.815
Income source - household-livestock	<b>U</b>	0.11163	0.09722	4.7		0.569
	<b>M</b>	0.10952	0.10901	0.2	96.4	0.987
Income source - household-business	<b>U</b>	0.06047	0.03935	9.7		0.230
	<b>M</b>	0.05714	0.07206	-6.8	29.3	0.575
Number of adults in household	<b>U</b>	3.2605	2.7315	42.2		0.000
	<b>M</b>	3.2238	3.2375	-1.1	97.4	0.913
Household members- no education	<b>U</b>	0.31628	0.442130	-26.1		*0.002
	<b>M</b>	0.32381	0.30843	3.2	87.8	0.735
Households toilet is pit latrine	<b>U</b>	0.83256	0.68981	33.9		0.000
	<b>M</b>	0.82857	0.8053	5.5	83.7	0.539
Household member own dwelling	<b>U</b>	0.9814	0.88657	38.9		0.000
	<b>M</b>	0.98095	0.97765	1.4	96.5	0.812
wealth_index_2009_q5_2	<b>U</b>	0.15814	0.21991	-15.8		0.064
	<b>M</b>	0.15814	0.15446	1.9	87.9	0.835
wealth_index_2009_q5_3	<b>U</b>	0.1619	0.19213	6.5		0.429
	<b>M</b>	0.2186	0.22309	-1.0	84.7	0.921
wealth_index_2009_q5_4	<b>U</b>	0.21905	0.15741	30.8		0.000
	<b>M</b>	0.28372	0.26077	4.9	84.0	0.643
wealth_index_2009_q5_5	<b>U</b>	0.28095	0.17361	19.0		0.020
	<b>M</b>	0.25116	0.25382	-1.5	92.0	0.884
Number of cows owned by household	<b>U</b>	0.24762	0.1.3889	2.8		0.742
	<b>M</b>	1.4558	1.6052	-5.5	-92.8	0.559

Number of Oxen owned by household	<b>U</b>	1.4762	0.87967	9.3		0.270
	<b>M</b>	1.0279	1.163	-7.5	19.0	0.453
Number of Sheep owned-household	<b>U</b>	1.0429	0.37731	10.6		0.193
	<b>M</b>	0.53023	0.52806	-1.0	91.0	0.924
Any member of household grew crops	<b>U</b>	0.51429	0.80787	28.6		0.001
	<b>M</b>	0.90698	0.88905	4.5	84.1	0.598
Any household land irrigated	<b>U</b>	0.90476	2.2315	22.6		0.008
	<b>M</b>	2.4651	2.4309	3.9	82.7	0.663
Household had agricultural land	<b>U</b>	2.4714	0.87106	10.5		0.175
	<b>M</b>	1.054	1.0092	2.0	80.5	0.840
Land area farmed	<b>U</b>	1.0448	1.2373	15.3		0.032
	<b>M</b>	1.7	1.4369	-1.1	93.0	0.814
Time to reach centre of community	<b>U</b>	21.619	31.634	-30.1		0.001
	<b>M</b>	21.957	21.615	1.0	96.6	0.853
<b>Observations</b>		<b>210</b>	<b>432</b>			

Sample	Ps R <sup>2</sup>	LR-chi2	p>Chi2	Mean bias	Median bias	B	R	%concern	%bad
<b>Unmatched</b>	0.179	147.24	0.000	24.5	24.4	104.8*	0.41*	50	32
<b>Matched</b>	0.005	3.09	1.000	2.8	2.0	17.2	1.15	23	5

\* If B>25%, R outside (0.5; 2)

# APPENDIX 3: ROBUSTNESS CHECKS

In order to address the validity of the results presented in Section 5, additional analyses with different estimation techniques were performed. This section presents the different econometric models used to test the robustness of the estimates presented in Section 5.

## 1 Multivariate regression

The first basic specification for estimating the impact of project participation is an Ordinary Least Squares (OLS) model (when the dependent is continuous) or probit model when the dependent is binary.

$$Y_i = \alpha + \beta_1 \text{Project participation}_i + \delta' X_i + \varepsilon_i$$

Where  $Y_i$  is the dependent variable;  $X_i$  is a vector of household characteristics also employed in the non-parsimonious model in table A1.1; finally the variable of interest is the dummy variable *Project Participation* that assumes value equal to one when the household is enrolled in the project, zero otherwise. When the dependent variable  $Y_i$  is not continuous but binary, we implemented probit models rather than Linear Probability Models estimated using OLS. It is important to note that in the absence of randomised allocation of the project among the population in our sample, OLS and probit models fail to identify the causal effect of the programme, and can only be used as additional check for our estimates. In the tables following only  $\beta_1$  will be reported.

## 2 Propensity Score Matching – Nearest Neighbour

With the Nearest Neighbour matching individual from the comparison group is chosen as a matching partner for a treated individual that is closest in terms of propensity score. Treated individuals can be match also with  $n$  nearest neighbours. The matching can be 'with replacement' when untreated individual can be used more than once, or 'without replacement' when it can be used only once. Allowing replacement, the average quality of the matching increase and the bias decrease, which is particularly useful when the distribution of the propensity score is very different between project participants and the control group (Caliendo and Kopeinig, 2008). Estimates in the analysis will allow matching with five nearest neighbours with replacement.

## 3 Propensity Score Matching – Caliper

In order to avoid the difference in propensity score for participants and nearest non-participant neighbour being too high, a caliper matching can be used. This procedure imposes a threshold ( $\delta > 0$ ) on the maximum propensity score distance. Estimates in this analysis will impose a caliper of 0.05.

## 4 Propensity Score Weighting

Following the example of Hirano and Imbens (2001),<sup>8</sup> we implemented a regression adjustment with weights based on the propensity score. The average treatment effect can be estimated in a parametric framework as follows:

$$Y_i = \alpha + \beta_1 \text{Project participation}_i + \delta_2' Z_i + \delta_1' X_i + \varepsilon_i$$

Where  $Y_i$  represents the outcome of interest;  $\text{Project participation}_i$  is a dummy binary variable equal to one if an individual/household is enrolled into the programme and zero otherwise;  $X_i$  is a vector of matching covariates used to estimate the propensity

score match; and  $Z_i$  is a vector of control variables which cannot be used for the matching as they are not supposed to influence project participation. The regression is estimated with weights equal to one for the treated units and  $\hat{e}(x)/(1 - \hat{e}(x))$  for control units.

This parametric regression analysis framework has the advantage to explore heterogeneity in the treatment effect. Moreover it allows controlling for variables that cannot be included in the propensity score equation. The robustness check tables will only report  $\beta_1$ .

**Table A3.1: Agricultural production**

	(1)	(2)	(3)	(4)
	OLS/probit	PSM NN	PSM Caliper	Propensity Score Weighting
<b>Total quantity of crops produced in the last 12 months (coffee, mango, maize, banana and chate) (Quintals)</b>	968.789** (479.713)	977.200* (580.700)	977.200* (559.417)	910.516* (482.994)
<b>N</b>	642	642	642	642
<b>Log – total quantity of crops produced in the last 12 months</b>	0.692*** (0.146)	0.876*** (0.245)	0.876*** (0.232)	0.638*** (0.165)
<b>N</b>	642	642	642	642

Robust standard errors in parentheses. PSM estimates bootstrapped 1,000 repetitions. \* p<0.1, \*\* p<0.05, \*\*\* p<0.01

**Table A3.2: Agricultural selling and revenues**

	(1)	(2)	(3)	(4)
	OLS/probit	PSM NN	PSM Caliper	Propensity Score Weighting
<b>Total quantity of crops sold in the last 12 months (Quintals) (coffee, mango, maize, banana and chate) Quintals</b>	1784.770*** (637.294)	1645.000*** (713.504)	1645.00** (717.143)	1726.988*** (612.804)
<b>N</b>	642	642	642	642
<b>Log – total quantity of crops in the last 12 months</b>	0.724*** (0.118)	0.713*** (0.185)	0.713*** (0.184)	0.642*** (0.131)
<b>N</b>	642	642	642	642
<b>Total revenue from sales of crops in the last 12 months (coffee, mango, maize, banana and chate) (BIR)</b>	4258.418*** (1415.059)	4899.471*** (1502.984)	4899.471*** (1564.984)	4060.139*** (1344.308)
<b>N</b>	642	642	642	642

Robust standard errors in parentheses. PSM estimates bootstrapped 1000 repetitions. \* p<0.1, \*\* p<0.05, \*\*\* p<0.01

**Table A3.3: Food consumption**

	(1)	(2)	(3)	(4)
	OLS/probit	PSM NN	PSM Caliper	Propensity Score Weighting
<b>Daily per capita food consumption</b>	0.939 (0.765)	1.115 (1.038)	1.115 (1.025)	1.108 (0.773)
<b>N</b>	642	642	642	642
<b>ln (daily per capita food consumption)</b>	0.128** (0.060)	0.142* (0.082)	0.142* (0.081)	0.138** (0.064)
<b>N</b>	642	642	642	642

Robust standard errors in parentheses. PSM estimates bootstrapped 1,000 repetitions. \* p<0.1, \*\* p<0.05, \*\*\* p<0.01

**Table A3.4: Quantity of coffee produced, sold and revenues**

	(1)	(2)	(3)	(4)
	OLS/probit	PSM NN	PSM Caliper	Propensity Score Weighting
<b>Quantity of coffee produced in the last 12 months (Quintals)</b>	583.728*** (222.466)	640.124*** (232.996)	640.124*** (223.177)	543.857*** (243.362)
<b>N</b>	642	642	642	642
<b>Quantity of coffee sold in the last 12 months (Quintals)</b>	1443.504*** (438.740)	1424.429*** (432.505)	1424.429*** (432.687)	1403.023*** (414.284)
<b>N</b>	642	642	642	642
<b>Revenues from coffee sales in the last 12 months (BIR)</b>	4850.968*** (1403.249)	5107.095*** (1442.594)	5107.095*** (1477.537)	4662.861*** (1336.533)
<b>N</b>	642	642	642	642

Robust standard errors in parentheses. PSM estimates bootstrapped 1,000 repetitions. \* p<0.1, \*\* p<0.05, \*\*\* p<0.01

**Table A3.5: Changes in coffee production, prices, income and processing machines**

	(1)	(2)	(3)	(4)
	OLS/probit	PSM NN	PSM Caliper	Propensity Score Weighting
<b>Increase in total household production of coffee since 2009</b> %	0.268*** (0.035)	0.138** (0.059)	0.138** (0.060)	0.190*** (0.042)
<b>N</b>	642	642	642	642
<b>Increase in household income from coffee sales since 2009</b> %	0.265*** (0.035)	0.152** (0.060)	0.152** (0.061)	0.172*** (0.042)
<b>N</b>	642	642	642	642
<b>Increase in coffee prices since 2009</b> %	0.215*** (0.032)	0.133** (0.053)	0.133** (0.054)	0.100*** (0.035)
<b>N</b>	642	642	642	642
<b>Increase in number of coffee processing machines since 2009</b> %	0.237*** (0.040)	0.237*** (0.041)	0.237*** (0.041)	0.241*** (0.048)
<b>N</b>	642	642	642	642

Robust standard errors in parentheses. PSM estimates bootstrapped 1,000 repetitions. \* p<0.1, \*\* p<0.05, \*\*\* p<0.01

**Table A3.6: Household income**

	(1)	(2)	(3)	(4)
	OLS/probit	PSM NN	PSM Caliper	Propensity Score Weighting
<b>Total daily per capita consumption</b>	5.428*** (1.491)	5.843*** (2.022)	5.843*** (1.921)	5.752*** (1.442)
<b>N</b>	642	642	642	642
<b>ln(total daily per capita consumption)</b>	0.247*** (0.059)	0.282*** (0.085)	0.282*** (0.082)	0.251*** (0.056)
<b>N</b>	642	642	642	642
<b>1[Global Indicator</b>	0.156*** (0.039)	0.219*** (0.065)	0.219*** (0.062)	0.218*** (0.048)
<b>N</b>	642	642	642	642

Robust standard errors in parentheses. PSM estimates bootstrapped 1,000 repetitions. \* p<0.1, \*\* p<0.05, \*\*\* p<0.01

**Table A3.7: Wealth index**

	1	2	3	4
	OLS/probit	PSM NN	PSM Caliper	Propensity Score Weighting
<i>Difference in wealth index between 2009 and date of survey</i>	0.029*** (0.011)	0.031** (0.014)	0.031** (0.015)	0.036*** (0.012)
<i>N</i>	642	642	642	642
<i>Number of quintiles of wealth index in which household increased</i>	0.035 (0.103)	0.100 (0.159)	0.100 (0.159)	0.176 (0.116)
<i>N</i>	642	642	642	642

Standard errors in parentheses; \* p<0.1, \*\* p<0.05, \*\*\* p<0.01; PSM estimates are bootstrapped with 1,000 repetitions, with standard errors clustered by community. All means are calculated *after* matching.

# NOTES

- 1 A mature project has been running for long enough – typically at least 2.5 years – to have a reasonable expectation of impact, with either an expenditure rate of at least 70% or completion of most project activities.
- 2 As well as calculating the results from the propensity-score matching kernel model, comparable estimates to those shown in Table 5.2 were calculated using an alternative propensity-score matching model (nearest neighbour with replacement) and with three different regression models. The results were similar in magnitude and significance to those shown in Table 5.4. Similarly, alternative models are used to check on the robustness of each of the results included in this report. Where the size or significance of the outcome estimates differ markedly between models, this is noted in the text.
- 3 See Gujarati, Damodar N. (2003) *Basic Econometrics*: Fourth Edition. New York: McGraw Hill.
- 4 Per capita figure refers to adult equivalent units. Daily total consumption was divided by a factor representing household size, to generate a per-day, per-person expenditure figure. To reflect the existence of economies of scale within households, and the lower consumption needs of children, the formula used for calculating household size is  $(A + \alpha K)^\theta$ , where A is number of adults in the household; K is the number of children;  $\alpha$  is the consumption of a child relative to an adult; and  $\theta$  stands for the extent of economies of scale. This Effectiveness Review follows the common practice of setting  $\alpha$  equal to 0.33 and  $\theta$  equal to 0.9.
- 5 When items are used in a scale or index, they should all measure the same underlying latent construct (e.g. household wealth status). The items, then, must be significantly correlated with one another. Cronbach's alpha is a measure of this inter-item correlation. The more the variables are correlated, the greater is the sum of the common variation they share. If all items are perfectly correlated, alpha would be 1 and 0 if they all were independent from one another. For comparing groups, an alpha of 0.7 or 0.8 is considered satisfactory. See: Bland, M. J. & Altman, D. G. 1997. Statistics notes: Cronbach's alpha. *BMJ*, 314, 572.
- 6 Caliendo, M. and Kopeinig, S. 2008. Some Practical Guidance for the Implementation of Propensity Score Matching, *Journal of Economic Surveys*, Wiley Blackwell, vol. 22(1), pages 31–72.
- 7 Bootstrapping is a statistical procedure where repeated samples are drawn from the original sample and parameters, such as standard errors, are re-estimated for each draw. The bootstrapped parameter is calculated as the average estimate over the total number of repeated draws.
- 8 Hirano, K. and Imbens G.W. (2001), Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services & Outcomes Research Methodology*, vol. 2, pp. 259–278.

## Oxfam Effectiveness Reviews

For more information, or to comment on this report, email [ppat@oxfam.org.uk](mailto:ppat@oxfam.org.uk)

© Oxfam GB November 2015

This publication is copyright but the text may be used free of charge for the purposes of advocacy, campaigning, education, and research, provided that the source is acknowledged in full. The copyright holder requests that all such use be registered with them for impact assessment purposes. For copying in any other circumstances, or for re-use in other publications, or for translation or adaptation, permission must be secured and a fee may be charged. E-mail [policyandpractice@oxfam.org.uk](mailto:policyandpractice@oxfam.org.uk).

The information in this publication is correct at the time of going to press.

Oxfam GB, Oxfam House, John Smith Drive, Cowley, Oxford, OX4 2JY, UK.

# OXFAM

Oxfam is an international confederation of 18 organisations networked together in over 90 countries, as part of a global movement for change, to build a future free from the injustice of poverty:

Oxfam America ([www.oxfamamerica.org](http://www.oxfamamerica.org))

Oxfam Australia ([www.oxfam.org.au](http://www.oxfam.org.au))

Oxfam-in-Belgium ([www.oxfamsol.be](http://www.oxfamsol.be))

Oxfam Brasil ([www.oxfam.org.br](http://www.oxfam.org.br))

Oxfam Canada ([www.oxfam.ca](http://www.oxfam.ca))

Oxfam France ([www.oxfamfrance.org](http://www.oxfamfrance.org))

Oxfam Germany ([www.oxfam.de](http://www.oxfam.de))

Oxfam GB ([www.oxfam.org.uk](http://www.oxfam.org.uk))

Oxfam Hong Kong ([www.oxfam.org.hk](http://www.oxfam.org.hk))

Oxfam India ([www.oxfamindia.org](http://www.oxfamindia.org)) Oxfam Intermón (Spain)

([www.oxfamintermon.org](http://www.oxfamintermon.org)) Oxfam Ireland ([www.oxfamireland.org](http://www.oxfamireland.org))

Oxfam Italy ([www.oxfamitalia.org](http://www.oxfamitalia.org))

Oxfam Japan ([www.oxfam.jp](http://www.oxfam.jp))

Oxfam Mexico ([www.oxfammexico.org](http://www.oxfammexico.org))

Oxfam New Zealand ([www.oxfam.org.nz](http://www.oxfam.org.nz))

Oxfam Novib ([www.oxfamnovib.nl](http://www.oxfamnovib.nl))

Oxfam Québec ([www.oxfam.qc.ca](http://www.oxfam.qc.ca))

Please write to any of the agencies for further information, or visit [www.oxfam.org](http://www.oxfam.org).

