



OXFAM

**Guatemala Highlands Value Chain
Development Alliance**

Project Effectiveness Review

Women's Empowerment Global Outcome Indicator



Oxfam GB

November 2012

Table of Contents

Executive Summary	4
1 Introduction and purpose	5
2 The Project: Guatemala Highlands Value Chain Development Alliance	6
3 Intervention logic of the support provided	7
4 Impact Assessment Design	8
4.1 Limitations in pursuing the gold standard.....	8
4.2 Alternative evaluation design pursued.....	9
4.2 Reconstruction of baseline.....	10
4.3 Selection of comparison group.....	10
5 Outcome indicators	11
5.1 Women’s empowerment global indicator.....	11
5.2 Other outcome measures.....	13
Self-reported income change	14
Ability to meet basic needs.....	14
Household income	14
Household food security.....	15
Household ownership of assets	15
Agricultural production and other productive activities	16
5.3 Measuring project exposure and adoption of improved production techniques	16
6 Methods of Data Collection and Analysis	17
6.1 Data collection	17
6.2 Data analysis.....	18
6.3 Main problems and constraints encountered.....	19
7 Results	20
7.1 General characteristics.....	20
7.2 Intervention exposure	22
7.3 Evidence of impact on outcome measures	23
7.3.1 Agricultural activities	23
7.3.2 Household income and wellbeing	28
7.3.3 Attitudes towards gender roles.....	33
7.3.4 Women’s influence in community-level producer associations.....	34
7.3.5 Women’s involvement in household decision-making.....	35
7.3.6 Women’s self-efficacy.....	37

7.3.7	Gender balance in household expenditure.....	38
8	Conclusion and Programme Learning Considerations	39
8.1	Conclusions.....	39
8.2	Programme Learning Considerations	40

Executive Summary

Under Oxfam Great Britain's (OGB) Global Performance Framework (GPF), samples of mature projects are being selected at random each year to undergo a rigorous assessment of their effectiveness. In the financial year of 2011/12, the Guatemala Highlands Value Chain Development Alliance (GUAB49) was selected for evaluation against OGB's global indicator for women's empowerment:

- **Proportion of supported women demonstrating greater involvement in household decision-making and influencing affairs at the community level.**

This project, together with two related projects (GUAB38 and GUAB62), have provided support to members of producer associations in the Department of Sololá, in partnership with the Asociación de Desarrollo Agrícola y Microempresarial (ADAM), and in alliance with Sysco, Superior Foods and Alimentos Sumar, the purchasers of the produce of the four associations. The primary objectives of these projects have been to improve household food security and strengthen agricultural livelihoods. Activities have included: agricultural extension support, distribution of fertiliser, training on the cultivation of kitchen gardens, provision of productive infrastructure to associations, and support in establishing linkages to private-sector vegetable exporters. At the same time, the activities are seeking to strengthen the position of women within the associations and within their households, with many of the activities carried out specifically with female members of the associations.

In February 2012, with the support of an external consultant, a team of enumerators administered a household survey to 383 members of producers' associations and producers' groups in the Department of Sololá. As well as members of the supported associations, enumerators interviewed members of comparable associations and groups that had not received similar external support. The survey was designed to allow evaluation of the projects against their intended outcomes and against OGB's global indicator for women's empowerment. At the analysis stage, the statistical tools of propensity score matching and multivariable regression were used to control for measured differences between the supported and comparison households.

The effectiveness review found evidence of positive impact on the ability of women to influence decisions in producer associations, as well as on attitudes to women's economic roles among women themselves. These effects are mostly concentrated on those women who were not directly members of the producer associations themselves, but who are wives or partners of male association members. To some extent, this may be because many of the women who are direct beneficiaries are members of women-only associations, and, consequently, there is less potential for their influence in their associations to change, as compared to women in mixed associations. There is no evidence to date that the projects have impacted women's influence in household level decision-making; the extent of women's involvement in such decision-making was found to be approximately the same among the supported and comparison households.

On the livelihoods outcomes, results are mixed. The projects appear to have been highly successful in encouraging the adoption of some improved agricultural techniques and cultivation of broccoli and other vegetables. However, there is, as of yet, no indication that this has led to higher returns to producer households. Surprisingly, the distribution of significant quantities of fertiliser in 2011 had no detectable effect on maize yields. Moreover, there is no evidence of an effect of activities carried out thus far on household income, food security or other measures of wellbeing.

Suggestions to enable to programme team to learn from this effectiveness review include:

- Consider whether and how project activities are intended to reinforce each other when designing projects, as well as how evaluation will be carried out.
- Consider using the sample interviewed for the project effectiveness review as a baseline for the full project (GUAB62).
- Investigate why the fertiliser distribution and other agricultural support did not lead to improved agricultural productivity or increased household income.

1 Introduction and purpose

Oxfam GB has developed a Global Performance Framework (GPF) as part of its effort to better understand and communicate its effectiveness and enhance learning across the organisation. This framework requires programme/project teams to annually report generic output data across six thematic indicator areas. In addition, a sample of sufficiently-mature projects (e.g. those closing during a given financial year) associated with each thematic indicator area are being randomly selected each year and rigorously evaluated. One key focus is on the extent they have promoted change in relation to relevant OGB global outcome indicators.

This report documents the findings of the project effectiveness review, focusing on outcomes related to women's empowerment.

The following global outcome indicator was endorsed for the livelihoods support thematic area:

- **Proportion of supported households demonstrating greater income, as defined by consumption and expenditure per day per capita**

The conceptual underpinnings of this indicator are presented in Section 3 below. The field work for the effectiveness review of the Guatemala Highland Value Chain Development Alliance (GUAB49), which took place in February 2012, was part of an effort to assess progress against this indicator.

This report presents the effectiveness review's findings. Section 2 provides brief background information on the projects and the context in which the support is being provided, while Section 3 explains the projects' intervention logic. Section 4, Section 5, and Section 6 follow by presenting the outcome measures used, the impact evaluation design pursued, and the methods of data collection and analysis, respectively. Section 7 is the longest section of this document. Its subsections present descriptive statistics, data on intervention exposure, and finally the overall differences between the targeted producers and the producers that were selected as comparators. Section 8 concludes with a summary of findings and some suggested learning considerations.

2 The Project: Guatemala Highlands Value Chain Development Alliance

The Guatemala Highlands Value Chain Development Alliance is one of a series of three projects carried out with a group of producer associations in the Department of Sololá.

The Guatemala Highlands Value Chain Development Alliance is a series of three projects which are intended to support small-scale producers in diversify and increase returns from their agricultural production. The beneficiaries are members of several producer associations located in the Department of Sololá, and the implementation is carried out by the Asociación de Desarrollo Agrícola y Microempresarial (ADAM).

The first of these projects, GUAB38, was launched to support food security and livelihoods in the aftermath of the devastation caused by Tropical Storm Agatha in May 2010. The project's interventions included supporting selected members of these associations to establish kitchen gardens and increase maize production (through the provision of technical support and fertiliser), as well as providing each association with three silos of maize for selling to members when prices are high. In three of the associations, ADAM also established "revolving fund" credit groups for a selection of members.

In 2011, the project GUAB49 constructed infrastructure for four of the associations, including composting plants, collection and storage centres for produce, and seedling banks. Also at this time, ADAM and Oxfam formed a partnership with a private-sector exporter to support producers to farm broccoli on a commercial scale. The newer project, GUAB62, is now building on this base to strengthen producers' linkages with commercial exporters. Moreover, under GUAB49, a small number of women from two of the associations were trained in the marketing of handicrafts and agricultural produce, using Oxfam's women's economic leadership (WEL) methodology.

The producer associations existed before the start of these projects, and ADAM began supporting them through a project funded by a different donor in 2008. ADAM has actively encouraged the participation of women in these associations, and many of the activities of the Oxfam-supported projects have specifically aimed at promoting greater involvement for women. For example, the committees formed to manage the maize silos provided by Oxfam and ADAM were composed solely of women. (One of the associations included in all three projects has only women members, but the other associations have mixed membership.)

Since the various projects have been carried out with the same group of producer associations – coupled by the fact that many of the same individual producers have benefited from more than one of the activities – the specific impact of each project cannot be isolated. More specifically, the four associations included in GUAB49 also participated in most of the interventions of GUAB38 (with the exception of the revolving funds activity, which was implemented solely in the three associations *not* subsequently included in GUAB49). This effectiveness review assessed the impact of the two projects GUAB38 and GUAB49, but only in the four producer associations which are included in GUAB49. There is also significant overlap with the beneficiaries from GUAB62, but the activities of this project only started around the time of the review's field work.

3 Intervention logic of the support provided

Figures 3.1 and 3.2 illustrate simple ‘theories of change’, showing the steps through which the project’s key interventions are intended to result in improvements in food security, livelihoods and women’s empowerment.

The two projects under consideration provided agricultural extension services and other forms of support aimed at agricultural production. Under the first of the three projects (GUAB38), a particular focus was on maize production, with many participants receiving a free distribution of fertiliser from the project. Starting in 2011, the GUAB49 project has instead focussed on the production of broccoli for sale to a private-sector exporter. All these activities were intended to improve productivity, in order to either produce more food for consumption within the household or generate income. Additionally, some women from each association were trained in producing vegetables, with the intention of improving diets, as well as potentially generating some extra household income.

Project activities include technical support for the production of value-chain products, as well as support to producers to bring their products to market.

Under the GUAB38 project, ADAM also provided three silos of maize to each association for onward sale to their members. The maize was intended to be sold at a reduced price during the season when prices are typically high, thereby, enabling households to reduce their spending on food.

The infrastructure investments made under project GUAB49 do not feature on this theory of change diagram because the infrastructure was not operational at the time of the survey and therefore cannot be expected to have brought benefits to the members. Note also that the “revolving fund” savings groups do not feature here because they were implemented (under GUAB38) in the three associations not targeted by GUAB49 and hence not covered by the effectiveness review.

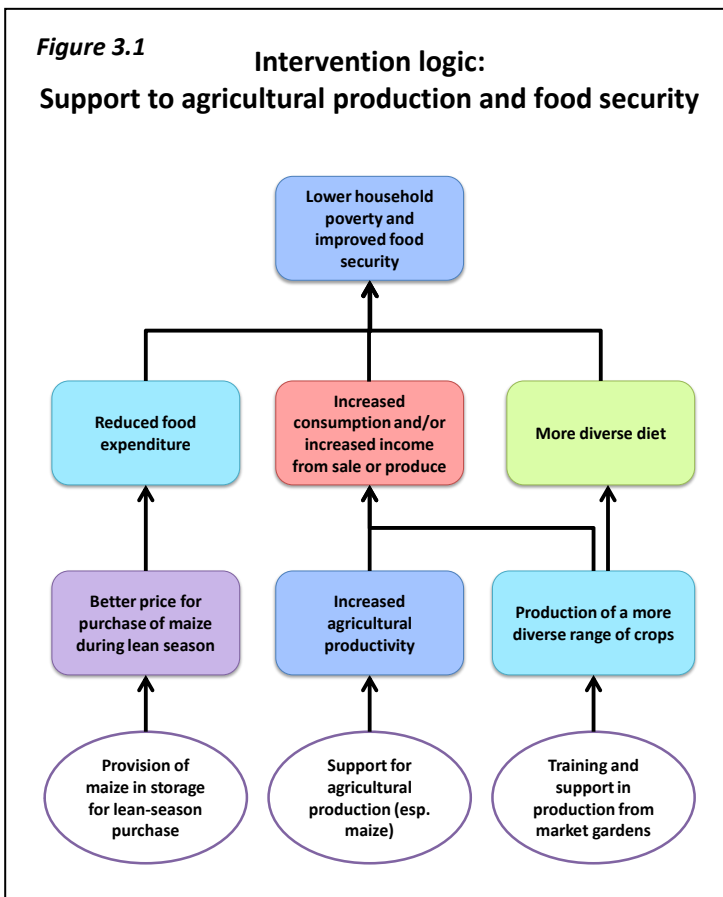
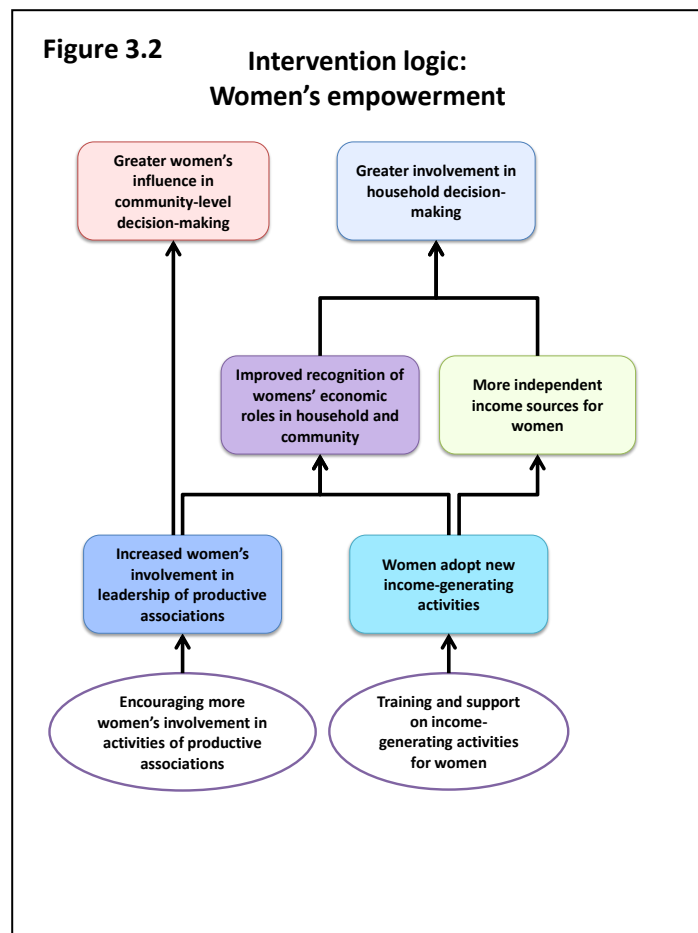


Figure 3.2 shows the channels through which these projects (and ADAM's support to the associations in the general) are expected to empower women. ADAM and Oxfam, for instance, have not only encouraged women to participate in the producer associations but also created conditions for them to take on positions of responsibility (such as in managing the donated maize and supervising the infrastructure investments). The provision of training in income-generating activities (e.g. kitchen gardens) was further intended to complement this, ultimately changing popular attitudes towards women, as well as elevating their status and influenced within their communities. The additional income was additionally expected to give them more decision-making influence within the household.



4 Impact Assessment Design

4.1 Limitations in pursuing the gold standard

The core challenge of a social impact evaluation is to credibly estimate the net effect of an intervention or programme on its participants. An intervention's net effect is typically defined as the average gain participants realise in outcome (e.g. income) from their participation. In other words:

Impact = average post-programme outcome of participants – what the average post-programme outcome of these same participants would have been had they never participated

This formula seems straightforward enough. However, *directly* obtaining data on the latter part of the equation – commonly referred to as the counterfactual – is logically impossible. This is because a person, household, community, etc. cannot *simultaneously* both participate and not participate in a programme. The counterfactual state of a programme's participants can therefore never be observed directly; it can only be estimated.

The randomised experiment is regarded by many as the most credible way of estimating the counterfactual, particularly when the number of units (e.g.

The aim of the effectiveness review is to estimate the difference that Oxfam GB's support has made to the supported producers.

people, households, or, in some cases, communities) that are being targeted is large. The random assignment of a sufficiently large number of such units to intervention and control groups should ensure that the statistical attributes of the two resulting groups are similar in terms of a) their pre-programme outcomes (e.g. both groups have the same average incomes); and b) their observed characteristics (e.g. education levels) and unobserved characteristics (e.g. motivation) that affect the outcome variables of interest. In other words, randomisation works to ensure that the *potential outcomes* of both groups are the same. As a result – provided that threats such differential attrition and intervention spill-over are minimal – any observed outcome differences observed at follow-up between the groups can be attributed to the programme.

However, implementing an ideal evaluation design like this is only possible if it is integrated into project design from the start, since it requires the introduction of some random element that influences participation. To evaluate an ongoing or completed programme – as in this project effectiveness review – or one where randomisation is judged to be impractical, it is therefore necessary to apply alternative techniques to estimate the counterfactual as rigorously as possible.

4.2 Alternative evaluation design pursued

There are several evaluation designs when the comparison group is non-equivalent that can – particularly when certain assumptions are made – identify reasonably precise intervention effect estimates. One solution is offered by matching: find units in an external comparison group that possess the same characteristics, e.g. ethnicity, age, and sex, as those of the intervention group and match them on these characteristics. If matching is done properly, the observed characteristics of the matched comparison group will be identical to those of the intervention group. The problem, however, with conventional matching methods is that with large numbers of characteristics on which to match, it is difficult to find comparators with similar combinations of characteristics for each of the units in the intervention group. The end result, typically, is that only a few units from the intervention and comparison groups get matched up, thereby, not only significantly reducing the size of the sample but also limiting the extent to which the findings can be generalised to all programme participants (referred to as the “curse of dimensionality” in the literature).

The evaluation design involved comparing the Oxfam supported producers with non-supported producers, while statistically controlling for observed differences between them.

Fortunately, matching on the basis of the propensity score – the conditional probability of being assigned to the programme group, given particular background variables or observed characteristics – offers a way out. The way propensity score matching (PSM) works is as follows: Units from both the intervention and comparison groups are pooled together. A statistical probability model is estimated, typically through logit or probit regression. This is used to estimate programme participation probabilities for all units in the pooled sample. Intervention and comparison units are then matched within certain ranges of their conditional probability scores. Tests are further carried out to assess whether the distributions of characteristics are similar in both groups after matching. If not, the matching bandwidth or calliper is repeatedly narrowed until the observed characteristics of the groups are statistically similar. Provided that a) the dataset in question is rich and of good quality; b) the groups possess many units with common characteristics (i.e. there is a large area of common support); and c) there are no unobserved

differences lurking among the groups, particularly those associated with the outcomes of interest, PSM can produce good intervention effect estimates.

Multivariable regression is another approach that is also used to control for measured differences between intervention and comparison groups. It operates differently from PSM in that it seeks to isolate the variation in the outcome variable explained by being in the intervention group *net of other explanatory variables* (key factors that explain variability in outcome) included the model. In this way, multivariable regression controls for measured differences between the intervention and comparison group. The validity of both PSM and multivariable regression are founded heavily on the “selection on observables” assumption, and therefore treatment effect estimates can be biased if there are unmeasured (or improperly measures) but relevant differences existing between the groups. Both PSM and multivariable regression were employed during data analysis and efforts were made to capture key explanatory variables believed to be relevant in terms of the assessed outcomes, including details about the composition of the household and their economic activities at baseline.

4.2 Reconstruction of baseline

Propensity-score matching or multivariate regression typically work more effectively when individual-level data on the situation of respondents at baseline are available, allowing for time-invariant differences between the groups to be controlled for. In the case of this project, baseline data were not available. Instead, the project effectiveness review attempted to reconstruct such data by asking respondents to recall their situation before the project commenced. In order to maximise the accuracy of the recalled data, respondents were asked to visualise their household's situation at the time of Tropical Storm Agatha in May 2010. It is important to note that they were only asked to recall information which they could reasonably be expected to recall with clarity, including the conditions of the homes, ownership of assets and livestock, and varieties of crops cultivated in 2010.

Baseline data were reconstructed by asking about the household's situation before the project began in 2010.

4.3 Selection of comparison group

A key factor in ensuring the validity of any non-randomised impact evaluation design is to employ an appropriate comparison group. This is particularly true for ex-post, cross-sectional designs. Comparators that differ in relation to the baseline status of the outcome variable(s) of interest and/or who are subjected to different external events and influences will likely result in misleading conclusions about programme impact. Identifying a plausible comparison group is therefore critically important in non-experimental impact evaluation and is, generally speaking, not an easy task.

For this effectiveness review, the selection of a suitable comparison group was complicated because the projects were implemented with members of local producer associations. It is very likely that those who have chosen to join an association are systematically different from those who have not, in terms of their productive activities, motivation, and other factors. Had producers not involved in the associations been selected for comparison purposes, it is very likely that this would have resulted in a biased comparison. For this reason, it was seen as important to use comparison producers that are currently members of an association or some other producer group similar to those supported by ADAM and Oxfam.

To this end, the effectiveness review team worked with ADAM and municipal authorities to identify associations and producer groups that have not received support from ADAM but are located in areas with similar characteristics.

Table 4.1 compares the associations supported through GUAB49 and these purposively matched comparison groups.

Table 4.1: Supported cooperatives and selection of comparison areas

Association	Municipality	Gender of members	Dominant ethno-linguistic group	Comparison
El Buen Sembrador	Santa Catarina Ixtahuacán	Mixed	K'iche'	Producers' groups in other communities in the Municipality of Santa Catarina Ixtahuacán and in the Municipality of Nahualá
ALANEL	Santa Catarina Ixtahuacán	Female	K'iche'	Women producers' groups in other communities in the Municipality of Santa Catarina Ixtahuacán
ADICOSO	Concepción	Mixed	Kaqchikel	Producers' groups in the Municipality of San José Chacayá
ASDIC	San Antonio Palopó	Mixed	Kaqchikel	Newer members of ADAX, another producers' association in the Municipality of San Antonio Palopó, and neighbouring women producers

5 Outcome indicators

5.1 Women's empowerment global indicator

The Oxfam GB global indicator for women's empowerment is a measure which is positive when a woman has *both* greater involvement in household decision-making *and* greater ability to influence affairs at the community level, in comparison with a typical (median) woman in the comparison group. The survey tools used to measure involvement in household decision-making and influence in community-level decision-making will be considered separately here.

- **Women's involvement in household decision-making**

This measure considers both the breadth and depth of women's involvement in household decision-making. Breadth is defined in terms of the number of decision-making areas in which women are involved (e.g. decisions around food preparation, personal travel, and family planning). The instrument comprises 25 decision-making areas. Depth, on the other hand, is defined in relation to the extent of involvement in each area, i.e. exclusive involvement, very strong involvement, joint involvement, some involvement, or no involvement.

Women are also asked a follow-up question when they express they either have only some involvement or no involvement in a particular decision-making area: *To what extent do you think you could be more involved in making decisions in this area if you really wanted to?* If they respond by stating that that they could be more involved if they wanted to, either to a medium or high extent, this is taken to mean that they would have the freedom to be more involved but have chosen not to for one reason or another. In other words, the follow-up question is included in the questionnaire given the fact that some women may choose not to participate in particular household decision-making areas. What is of particular interest is the extent that women have the potential of being involved in the various decision-making areas. As such, based on their responses to the follow-up question, the women's decision-making score for the area in question is adjusted accordingly (e.g. kept the same if the women reports that she could not be involved in such decisions or is adjusted upwards if she reports that she could be more involved).

Both the breadth and depth of women's involvement in household decision-making was assessed.

- **Women’s influence in community-level decision-making**

The second component of the global indicator examines the extent to which women perceive they are able to influence the running of affairs at the community level. Female respondents are asked the extent they agree or disagree with these statements:

- There are opportunities for women to participate in producer associations in the community.
- If you wanted to give your opinion in association meetings, people would give you the opportunity to do so.
- There exist real opportunities to obtain a leadership position in these associations if you wanted to.
- It would be very difficult for you to influence how leaders are chosen.
- It would be quite easy for you to influence important decisions in these associations.
- There are large barriers to you holding leadership positions in these associations.
- If leaders were doing things you did not agree with, you would just have to adapt and could not do much to stop them.
- Things have really changed in these associations: now there are opportunities for women to influence how the associations are run.
- There are many initiatives happening in your community where your voice could never be heard in any meaningful way.

Women were also asked about their ability to influence decisions taken in community-level producer associations.

(In this effectiveness review, these questions were phrased so that they do not only apply if the respondent herself is a member of a community-level association: this is important since, as discussion in Section 6.1, wives and partners of male association members were also interviewed.)

As is apparent, some of these statements are presented in a positive sense and some in a negative sense. In the process of data analysis, the responses to the negative phrases were inverted, with points awarded according to the extent of agreement or disagreement with each phrase. Rather than simply using the raw scores as the basis of the measure, principal factor analysis was carried out on the nine items. This technique focuses on the variation in the data that is common in the responses, so reducing the amount of “noise” in the data. This increases precision in comparing group differences in attitudes.

- **Attitudes to women’s economic roles**

Female respondents were also asked to state the extent of their agreement or disagreement with each of these statements:

- The only really satisfying role for a woman is as a wife and mother.
- Women are as important as men in ensuring that the basic material needs of families are met.
- Girls should be encouraged to be ambitious in terms becoming economically independent when they reach womanhood.
- Women are not suited for work of great stress and responsibility.
- Women’s livelihood work is equally as important as their domestic work.
- A man should be responsible for providing money for his wife’s personal use even if she is capable of earning it herself.

Other indicators of empowerment collected in the survey were a measure of attitudes to women's economic roles and a measure of women's self-efficacy.

- Women's most important job is to look after the comforts of men and children.
- Households in our community would be much poorer if women stopped doing livelihood work.
- A situation where a woman spends the majority of her day away from the home to make money is not right.
- If a child falls ill, it is the mother's duty rather than the father's to take time away from productive activities to look after him or her.
- The saying "a woman's place is to take care of the home" is generally correct.
- A woman can be a good wife and mother even if she is involved in demanding livelihood activities.
- Women should worry less about their rights and more about becoming good wives and mothers.
- In general, women are equally capable of contributing to economic well-being than are men.
- If a woman gets too involved in livelihood activities, her family will likely suffer.

As in the case of the community-level influence questions, the responses to negative statements were inverted at the analysis stage and an overall index was created for each respondent's attitudes using factor analysis.

- **Women's self efficacy**

Finally, women's self-efficacy was measured using an adapted version of the General Self-Efficacy Scale (GSE). This is a four-point Likert scale that asks respondents the extent to which each of the following statements is true for them:

1. You can always manage to solve difficult problems if you try hard enough.
2. If someone opposes you, you can find the means and ways to get what you want.
3. It is easy for you to stick to your aims and accomplish your goals.
4. You are confident that you could deal efficiently with unexpected events.
5. Thanks to your resourcefulness, you know how to handle unforeseen situations.
6. You can solve most problems if you invest the necessary effort.
7. You can remain calm when facing difficulties because you can rely on your coping abilities.
8. When you are confronted with a problem, you can usually find several solutions.

A score for each respondent's self-efficacy was constructed, again using factor analysis.

5.2 Other outcome measures

As reviewed in Section 3 above, the support provided to the targeted households is intended to bring about a number of other outcomes – improved livelihoods – in addition to having an impact on women's empowerment. To evaluate the success on these other dimensions, data were therefore collected on a number of additional outcome measures.

Self-reported income change

Respondents were asked to make a judgement whether their overall income had increased, remained the same or decreased since 2010.

Ability to meet basic needs

Respondents were presented with the following four descriptions of household economic situations, and asked which matched their own situation most closely:

- Doing well: able to meet household needs by your own efforts, and making some extra for stores, savings, and investment.
- Breaking even: Able to meet household needs but with nothing extra to save or invest.
- Struggling: Managing to meet household needs, but depleting productive assets and/or sometimes receiving support.
- Unable to meet household needs by your own efforts: dependent on support from relatives living outside of your household or the community, government and/or some other organisation – could not survive without this outside support.

As well as the women's empowerment measures, the survey also collected data on various livelihood measures.

Household income

Measuring household wealth or socioeconomic position in low income countries is not straightforward, particularly in rural areas where respondents tend to be self-employed. Self-reported measures of total income are unreliable, given the wide variety of endeavours such populations engage in to generate income.¹ However, given that there is a widely recognised and strong association between household income and consumption,² one popular proxy measure used by the World Bank and other international institutions involves the aggregation of both household consumption and expenditure data.³ To capture data on this indicator, a household survey is administered that contains a consumption and expenditure module. The respondents are asked what types of food they consumed over the previous seven day period, as well as the particular quantity. The quantity is transformed into a monetary value, i.e. either how much they paid for the food item in question or, if the food item was from their own production, how much they would have paid if it was bought from the local market. The respondents are also asked how much they spent on particular regular non-food items and services from a list such as soap, toothpaste, and minibus fares over the past four weeks. Finally, they are asked for any household expenditure on non-regular non-food items such as school and hospital fees, clothes, and home repair over the last 12 months. For non-food items that are gender divisible, data are collected in a gender disaggregated fashion, thereby, enabling intra-household consumption inequality to be measured as well.

The household expenditure measure is calculated by converting each of the expenditure types into a per-day figure and adding them together. While dividing the above equation by household size as the overall denominator is recommended in the literature, using a more nuanced calculation is deemed important to avoid underestimating the wealth status of larger sized households relative to their smaller counterparts. The formula used for calculating household size is

¹ Morris, Saul, Calogero Carletto, John Hoddinott, and Luc J. M. Christianensen. (1999) *Validity of Rapid Estimates of Household Wealth and Income for Health Surveys in Rural Africa: FCND Discussion Paper No. 72*. Washington: International Food Policy Research Institute.

² See Gujarati, Damodar N. (2003) *Basic Econometrics: Fourth Edition*. New York: McGraw Hill.

³ Deaton, A and S. Zaidi. 2002. "Guidelines for constructing consumption aggregates for welfare analysis," Working Paper No. 135. The World Bank, Washington, D.C.

where A is number of adults in the household; K is the number of children; α is the consumption of a child relative to an adult; and β stands for the extent of economies of scale. This evaluation follows the common practice of setting α equal to 0.33 and β equal to 0.9,⁴ but the findings are not sensitive to reasonable changes in these parameters.

The expenditure variable is normally converted to a logarithmic scale, to improve the model fit in regression analysis and reduce the influence of outliers. The resulting variable can remain continuous and the average per capita consumption and expenditure can be calculated for the sample in question. It can also be transformed into a binary variable, so that the proportion of households living above a certain monetary figure can be calculated. For the Oxfam GB global indicator for livelihoods, the median expenditure level of the comparison group is used as the benchmark for creating the binary variable.

Household expenditure data were complemented by data on food security, asset wealth, and the household's judgement of its ability to meet basic needs.

Household food security

Household food security was measured using six questions adapted from the Household Food Insecurity Access Scale (HFIAS) developed by USAID's Food and Nutrition Technical Assistance (FANTA) Programme.⁵ Respondents were asked whether any of the following were true for them or other members of their household in the four weeks before the date of the survey:

- Did you or any household member have to eat some foods that you really did not want to eat because of a lack of resources to obtain other types of food?
- Did you or any household member have to eat a smaller meal than you felt you needed because there was not enough food?
- Did you or any household member have to eat fewer meals in a day because there was not enough food?
- Was there ever no food to eat of any kind in your house because of lack of resources to get food?
- Did you or any household member go to sleep at night hungry because there was not enough food?
- Did you or any household member go a whole day and night without eating anything because there was not enough food?

For each question which was answered positively, the respondent was then asked how frequently this situation occurred during the four weeks. A score was generated based on the frequency of these events.

Household ownership of assets

Household consumption and food security provide good indications of the household's current economic situation, but, particularly in low-income contexts, they tend to be influenced strongly by current or very recent income patterns. These measures may not, therefore, fully reflect any long-term economic benefits from participation in a project. In order to assess the project's impact on more established household wealth status, the survey also asked the respondents about their ownership of livestock, household assets, and the condition of their homes. The full list of assets and other wealth indicators collected in the survey is shown in Table 5.1. Respondents were asked about

⁴ Ibid.

⁵ http://www.fantaproject.org/publications/hfias_intro.shtml

their ownership of these assets both at the time of the survey and in 2010, the baseline period.

Table 5.1: List of assets and other wealth indicators used to derive asset index

Livestock	Agricultural equipment	Household goods	Vehicles
Cattle	Hoe	Clock	Bicycle
Sheep	Machete	Table	Motorcycle
Goats	Axe	Bed with mattress	Car or motor vehicle
Pigs	Spade	Lamp (electric or gas)	Kayak
Donkeys	Rake	Iron (electric or coal)	
Horses	Pick	Sewing machine	Condition of house
Turkeys	Grass chopper	Grinding stone	Number of rooms
Chickens	Scythe	Cooking plate	Number of floors
Ducks	Chainsaw	Gas stove	Material used for walls
Rabbits	Cart	Electric stove	Material used for roof
	Silo	Microwave	Material used for floor
Property	Milking bucket	Fridge	Type of toilet
Ownership of house	Thermal milk flask	Blender	Electricity connection
	Sprayer	Mobile phone	
		Radio	
		CD player	
		Sound system	
		Television	
		DVD player	
		Fan	
		Generator	
		Solar panel	

Principal component analysis (PCA) was used to create a weighted index of asset ownership for 2010, as well another index of changes in asset ownership since 2010. PCA is a data reduction technique that narrows in on the variation in household asset ownership, assigning greater weight to those observations better off in those assets associated with this variation. This enables the relative wealth status of the households to be compared.

Agricultural production and other productive activities

Respondents were asked to provide details for their production and sales of agricultural products in the 12 months prior to the survey. In order to provide a fuller picture of the diversification of income sources within their household, respondents were also asked about the economic activities which each member of their household engages in. They were further asked to estimate the proportionate contribution of each activity to total household income, both currently and in 2010. This was facilitated by showing respondents a sheet with images of various sources of income and asking them to allocate 20 stones between the sources according to their situation.

5.3 Measuring project exposure and adoption of improved production techniques

To assess progress along the steps in the intervention logic models described in Section 3, it was necessary also to measure the extent to which the

respondents were exposed to different types of support targeted at the households. As such, the respondents were asked whether they had received specific forms of support during the 12 months prior to the survey, as well as whether each of these forms of support had been provided by a local producers' association or by some other (governmental or non-governmental) organisation. They were also asked which of the techniques recommended by ADAM for improved agricultural production they had implemented during the past year.

6 Methods of Data Collection and Analysis

6.1 Data collection

The effectiveness review team designed a household questionnaire to capture data on both the outcome variables presented in Section 5 above, as well as other key characteristics of the intervention and comparison producers. This questionnaire was tested in pilot communities during the training workshop for the enumerators and subsequently revised. Sixteen potential enumerators participated in the two-day training workshop. Based on their performance in this exercise and on their fluency in the appropriate languages, 11 enumerators were selected to carry out the field work. One more enumerator, who had been held in reserve after the training, was added to the team half-way through the field work.

The questionnaire was piloted during the enumerator training workshop.

It was clear that the specific interventions carried out under GUAB38 and GUAB49 had targeted specific members of the producer associations, rather than the membership of each association as a whole. Most of the associations have large membership lists, comprising of many members who have not directly participated in these projects. However, ADAM was able to provide the effectiveness review team with lists of individuals who had participated in each of the activities carried out under projects GUAB38 and GUAB49. These lists were consolidated to form a single list of 262 beneficiaries (142 women and 120 men). In each association, at least two thirds of the beneficiaries had participated in only one of the interventions from the two projects.

From this single list of beneficiaries, a sample of 150 were chosen at random to be interviewed. The random sample was carried out separately in each association and separately for women and men. This was to ensure that the balance in the survey sample would reflect the balance in the beneficiary group as a whole.⁶

Since the unmatched comparison data are either dropped or given less weight in PSM, it is better for the sample size of the comparison group to be proportionally larger to the intervention group. To that end, the field staff were given a target of 224 comparison interviews to conduct, split between women and men and between the geographic areas in the same proportion as the project beneficiaries. As described in Section 4.3, comparison respondents were selected from among the members of local producer associations or groups in each comparison area. In most cases, all the appropriate members

⁶ In fact, because of complications over the beneficiary lists in some cooperatives, the sample was slightly revised on the first day, which has led to women being over-sampled in the final dataset: women make up approximately 54 per cent of the beneficiary group, but 62 per cent of the final sample. For simplicity, the results quoted in the "overall" columns in Section 7 do not account for this slightly unequal weighting for males and females, but it is not thought that this significantly affects the results.

of each association or group were surveyed, so there was no selection mechanism. In the event, the targets for the numbers of comparison respondents were slightly exceeded.

The women's questionnaire was carried out with the wife or partner of the project participant, in cases where the participant was male.

The questionnaire was divided into two sections. The first section asked about the household's composition, productive activities, and expenditure, and could be responded to by any adult member of the household (though in practice the respondent was usually the direct beneficiary named on the lists provided by ADAM, sometimes accompanied by his or her spouse or partner). The second section was carried out specifically with women and was conducted out of hearing range of other household members. In the case where the direct project beneficiary was female, the women's questionnaire was administered to her. In cases where the direct beneficiary was male, the women's questionnaire was administered to his wife or partner, to assess whether there had been any indirect impact of the projects on her situation or attitudes. The results of gender outcomes in Sections 7.3.3 to 7.3.6 is disaggregated into results on female direct beneficiaries and on the wives and partners of male direct beneficiaries.

6.2 Data analysis

OGB created a data entry interface in Adobe Acrobat Pro, and the Consultant recruited and supervised data entry clerks to enter the data. The data were imported into Stata for analysis, the results of which are presented in the following sections. The analyses involved group mean comparisons using *t*-tests, propensity score matching (PSM) with Stata's *psmatch2* module, and various regression approaches. Kernel and nearest neighbour matching without replacement were the main methods used in implementing PSM. Variables used in the matching process were identified by first using backwards stepwise regression to identifying those variables that are correlated with the outcome measure of interest at a *p*-value of less than 0.25 (or 0.30 for smaller samples). The short-listed variables were then put into another stepwise regression model to identify those that are correlated with being a member of the intervention group. Covariate balance was checked following the implementation of each matching procedure. When covariate imbalance at *p*-values of 0.25 or less was identified, the bandwidth or calliper was reduced and the PSM procedure and covariate balance test implemented again. This was continued until all covariates were balanced at *p*-values greater than 0.25. Boot-strapped standard errors enabled the generation of confidence intervals to assess the statistical significance of the effect sizes. All the covariates presented in Table 7.1 below were included in the various regression approaches undertaken, i.e. regression with robust standard errors, robust regression (to reduce the influence of outliers), and regression with control functions (to attempt to control for unobserved differences between the intervention and comparison groups).

6.3 Main problems and constraints encountered

Three difficulties encountered in the course of the field work provide challenges in analysing the data:

The complexity of the project design and the partially overlapping groups of beneficiaries presented challenges for analysis.

- *Low maturity of some project activities.* Although the GUAB49 project had closed by the time of the effectiveness review, the construction of the infrastructure facilities (the main activity under that project) had only recently been completed and these facilities were not yet in operation. This effectiveness review could not, therefore, assess the impact of those facilities on the livelihoods of the beneficiaries. The activities assessed by the effectiveness review are, consequently, almost exclusively related to the activities carried out under project GUAB38. However, these activities were implemented less than 18 months prior to the effectiveness review, a period of time perhaps too short for the potential impacts of these activities to fully manifest. Of course, it is still possible that ADAM and Oxfam's efforts to involve women in the management of activities under both projects may have already had on women's, even if not on the livelihoods of the association members.
- *Multiple interventions with sub-groups of beneficiaries.* The two projects comprised of a variety of activities, each of them involving groups of members from multiple associations. There was little overlap between these groups. For instance, among the interviewed beneficiaries, the largest pairwise correlation in activities they participated in is 0.17, between the group involved in managing the maize silos and the group involved in the kitchen garden activity. This means that two-thirds of the beneficiaries interviewed participated in only one project activity. Hence, although the effectiveness review attempted to identify overall impacts of the project on the beneficiaries, there is little consistency across the beneficiary group in the degree or type of support they received from the projects. In addition, the small number of beneficiaries who participated in the women's economic leadership training means that it is not possible to derive specific estimates of impact from this activity.
- *Complication of working in multiple languages.* A significant practical consideration was that the project beneficiaries are speakers of two different indigenous languages, neither of which were spoken by the Oxfam staff or the consultant leading the field work. This presented challenges in recruiting and training enumerators who had native-level fluency in those languages. For the women's questionnaire, in particular, it was very important that the statements in the survey which the interviewees were asked to respond to were interpreted consistently across interviews and consistently between languages. The training was extended for a third day to allow this clarification and consistency checking to be done, and the two ADAM field staff provided an excellent level of support in this.

7 Results

7.1 General characteristics

Table 7.1 presents descriptive statistics of the interviewed households from both the intervention and comparison groups. The stars beside the number indicate differences between the two groups that are statistically significant at a 90 percent confidence level or greater.

Overall, the intervention and comparison groups are well balanced. There are only a few differences between the groups:

- Supported households are more than twice as likely to have some member who has formal employment than are comparison households.
- When the primary respondent was male, the women who responded to the women's questionnaire were significantly more likely to be the wife of the household head (as opposed to a daughter or some other relation) among the support households than the comparison households.
- Also among male respondents, the heads of supported households are approximately four years older on average than heads of comparison households.
- The male respondents live on average further from the nearest public transport stop than do male comparison respondents (11 minutes' walk, compared to 8 minutes' in the case of the comparison respondents). This difference is the opposite in the case of female respondents.

The supported households and comparison households were well matched, with few significant differences between them.

The last two of these factors are unlikely to be large enough in magnitude to have a great affect on the outcome variables to be considered, but the first two factors may very well be. In any case, all these differences in characteristics will be controlled in the data analysis process, the results for which are presented in Section 7.3.

Table 7.1 also demonstrates that a larger number of household members are members of a producers' association among the supported households than the comparison households. This is to be expected, and it is likely to be a consequence of the projects or, at least, reflect the better organisation of the supported associations than the comparison associations.

**Table 7.1:
Descriptive statistics for intervention and comparison respondents**

	Intervention mean	Compariso n mean	Overall		Female respondents		Male respondents	
			Difference	t-statistic	Differenc e	t-statistic	Difference	t-statistic
Household size	6.486	6.089	0.397	1.39	0.575	1.60	0.071	0.15
Number of adults	3.534	3.268	0.266	1.42	0.318	1.36	0.167	0.54
Number of children	2.953	2.821	0.131	0.64	0.257	1.02	-0.096	-0.28
Number of productive adults	3.230	2.991	0.238	1.39	0.304	1.43	0.118	0.41
Number of unproductive adults	0.304	0.277	0.027	0.36	0.014	0.15	0.049	0.36
Household head female	0.129	0.090	0.040	1.22	0.092*	1.94	-0.047	-1.63
Household head > 60 years old	0.142	0.128	0.014	0.40	-0.037	-0.78	0.103**	2.02
Age of household head	44.777	42.991	1.786	1.25	0.268	0.14	4.361**	1.98
Only one adult in household	0.027	0.034	-0.007	-0.38	0.003	0.12	-0.023	-1.15
All household members > 60 years old	0.027	0.021	0.006	0.36	0.002	0.08	0.012	0.44
HH head is of the dominant ethnic group in the area	0.973	0.970	0.003	0.16	0.012	0.53	-0.012	-0.44
HH head has some primary education	0.277	0.261	0.016	0.35	0.048	0.80	-0.035	-0.47
HH head has some secondary education	0.000	0.009	-0.009	-1.13	-0.007	-0.79	-0.012	-0.81
HH head is engaged in some productive activity	0.986	0.983	0.004	0.27	0.005	0.26	0.000	.
Some HH member has formal employment	0.176	0.081	0.095***	2.83	0.076*	1.79	0.127**	2.30
Some female HH member has formal employment	0.061	0.026	0.035*	1.74	0.010	0.39	0.078**	2.27
WQ respondent is wife of HH head	0.800	0.796	0.004	0.08	-0.063	-1.06	0.108**	2.25
WQ respondent is daughter of HH head	0.076	0.106	-0.030	-0.97	-0.013	-0.29	-0.057	-1.46
Age of WQ respondent	40.483	38.445	2.038	1.49	1.146	0.66	3.602	1.62
WQ respondent > 60 years old	0.095	0.079	0.017	0.56	-0.018	-0.48	0.077	1.62
W Q respondent is married	0.850	0.890	-0.040	-1.14	-0.090*	-1.77	0.037	1.45
WQ respondent is of the dominant ethnic group in the area	0.966	0.965	0.001	0.05	0.019	0.78	-0.030	-0.89
WQ respondent has some primary education	0.272	0.241	0.031	0.67	0.065	1.10	-0.025	-0.35
WQ respondent has some secondary education	0.000	0.009	-0.009	-1.14	-0.014	-1.12	0.000	.
WQ respondent is currently in education	0.007	0.017	-0.011	-0.88	-0.009	-0.54	-0.013	-0.83
WQ respondent is engaged in some productive activity	0.986	0.991	-0.005	-0.46	-0.004	-0.34	-0.006	-0.29
Number of members in the household of a producers' association	1.291	1.051	0.239***	4.29	0.171**	2.50	0.348***	3.76
Asset index 2010	0.209	-0.125	0.334	1.21	0.152	0.45	0.631	1.35
Asset poorest third in 2010	0.311	0.352	-0.041	-0.80	-0.074	-1.14	0.013	0.15
Asset middle third in 2010	0.341	0.326	0.015	0.29	0.009	0.13	0.025	0.30
Asset wealthiest third in 2010	0.348	0.322	0.027	0.52	0.066	1.01	-0.038	-0.45
Distance from house to nearest transport stop (minutes on foot)	8.919	9.060	-0.141	-0.16	-2.044*	-1.74	3.058**	2.34
Distance from community to municipal capital (minutes in a vehicle)	34.554	38.809	-4.254	-1.35	-3.570	-0.77	-5.265	-1.61
Acres farmed in 2010	8.079	7.168	0.911	0.79	0.881	0.56	0.851	0.56
Total acres owned in 2010	4.742	4.529	0.212	0.33	0.328	0.62	-0.049	-0.03
Observations	148	235	383		241		142	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

HH: household

WQ: women's questionnaire

7.2 Intervention exposure

The respondents were asked a number of questions about the support they received from their cooperative, as well as other organisations, in the previous three years. Figure 7.1 and Table 7.2 show the proportion of supported and comparison producers who reported having received the various types of support.

Many more of the supported households reported receiving various forms of support during 2011 than the comparison households.

It is clear that a higher proportion of the supported producers reported receiving the types of support provided by the projects (including support to maize production, distribution of fertiliser, and training for women on business planning). Also, as expected, there were no significant differences between the supported and comparison households with respect to types of support which the projects did not deliver, including the distribution of tools, food or cash. There is further no difference in their access to compost. This is, again, to be expected, given that the compost facilities constructed under project GUAB49 had not fully started production at the time of the survey.

It should be noted that these figures only reveal the proportion of members who had access to each service; they do not provide information on the intensity of the provision, or the quality, of each service.

Figure 7.1: Proportion of surveyed households receiving support from external organisations

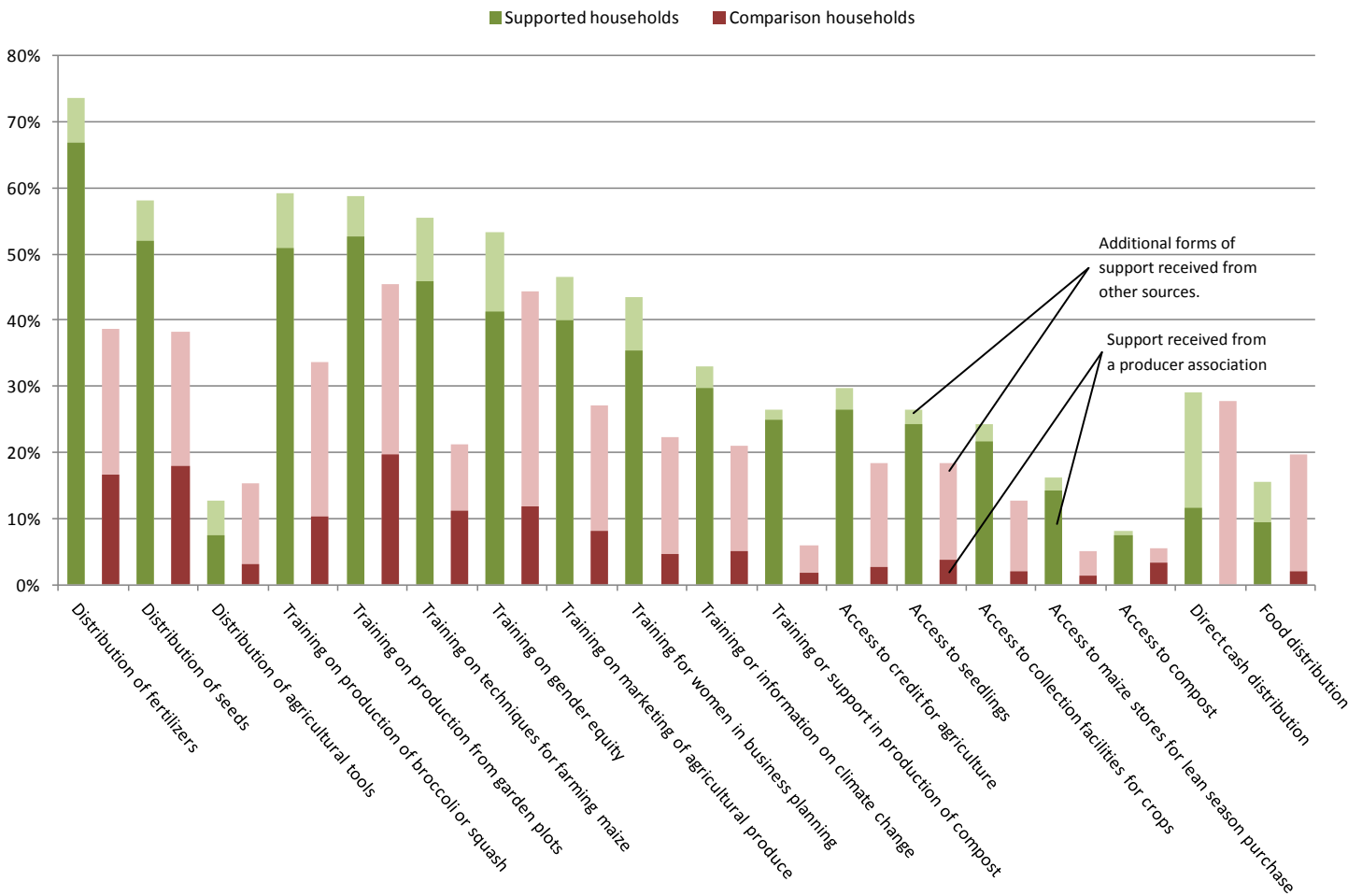


Table 7.2: Differences in support received (from all sources) in the 12 months previous to the survey

	Intervention Comparison		Overall		Female respondents		Male respondents	
	mean	n	Difference	t-statistic	Difference	t-statistic	Difference	t-statistic
Training on techniques for farming maize	0.554	0.213	0.341***	7.29	0.357***	6.06	0.315***	4.05
Training on production from garden plots	0.588	0.455	0.133**	2.54	0.203***	3.16	0.021	0.25
Training on production of broccoli/ squash	0.592	0.336	0.256***	5.05	0.253***	3.92	0.263***	3.22
Distribution of agricultural tools	0.128	0.153	-0.025	-0.67	0.006	0.12	-0.074	-1.45
Distribution of seeds	0.581	0.383	0.198***	3.85	0.204***	3.18	0.184**	2.17
Distribution of fertilisers	0.736	0.387	0.349***	7.07	0.311***	4.91	0.409***	5.27
Access to credit for agriculture	0.297	0.183	0.114***	2.62	0.132***	2.76	0.079	0.96
Access to collection facilities for crops	0.243	0.128	0.116***	2.94	0.066*	1.85	0.191**	2.39
Access to seedlings	0.264	0.183	0.081*	1.87	0.095*	1.72	0.058	0.84
Access to compost	0.081	0.055	0.026	0.99	0.040	1.25	0.002	0.04
Access to maize stores for lean season	0.162	0.051	0.111***	3.68	0.123***	3.35	0.091*	1.73
Training or information on climate change	0.331	0.209	0.123***	2.70	0.182***	3.19	0.024	0.31
Training on marketing of agri. produce	0.466	0.272	0.194***	3.95	0.242***	4.07	0.110	1.30
Training for women in business planning	0.435	0.222	0.213***	4.51	0.273***	4.55	0.116	1.52
Training on gender equity	0.534	0.443	0.091*	1.74	0.058	0.89	0.155*	1.95
Training or support in compost production	0.264	0.060	0.204***	5.86	0.246***	5.66	0.133**	2.30
Direct cash distribution	0.291	0.278	0.013	0.27	0.067	1.13	-0.079	-1.01
Food distribution	0.155	0.196	-0.040	-1.00	0.037	0.76	-0.172**	-2.49
Observations	148	235	383		241		142	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

7.3 Evidence of impact on outcome measures

7.3.1 Agricultural activities

Since it is clear that the intended beneficiaries of the projects received more external support than their comparators, the next stage is to examine the extent to which there is evidence that this support resulted in positive impacts. We will start by investigating the projects' impact on household agricultural production.

Much of the technical support provided by ADAM intended to encourage producers to adopt various improved production techniques. The survey asked about the adoption of five techniques: use of organic fertiliser, use of improved seeds, disease control, crop rotation, and other soil conservation techniques. An overall score was constructed for how many of these techniques each respondent had used in the year previous to the survey, with the techniques weighted by principal component analysis to highlight those with greatest variation. As can be seen in Table 7.3, adoption of the improved agricultural techniques was clearly higher among the supported producers than among the comparison producers. Examining the disaggregated data for the different techniques (not shown here) shows significantly higher adoption among the supported producers of improved seeds and disease control methods, and perhaps of crop rotation, but not of green manure or of other soil conservation methods.

Table 7.3: Number of improved agricultural techniques applied by the household during 2011

<i>Unadjusted:</i>	
Sample mean	2.314
Intervention mean	2.707
Comparison mean	2.065
Unadjusted difference	0.643*** (4.04)
Observations:	379
<i>PSM (ATT)</i>	
Post-matching difference (kernel)	0.572*** (3.55)
Observations:	367
Post-matching difference (no replacement)	0.416** (2.42)
Observations:	367
<i>Multivariable Regression:</i>	
MVR coefficient (fixed effects; robust standard errors)	0.586*** (3.77)
Observations:	341
MVR coefficient (robust regression)	0.586*** (3.77)
Observations:	341
MVR coefficient with control functions (robust SE)	0.580*** (3.72)
Observations:	339

t statistics in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

Supported households reported applying significantly more of the improved agricultural techniques than comparison

One of the main interventions was the provision of fertiliser and technical support for the production of maize. Data were collected on the specific quantities of maize produced and area cultivated by each household during 2011, thereby enabling the calculation of the yield per acreage. These figures are shown in the first three columns of Table 7.4. Maize is produced mostly for household consumption, but an indicator of wealth is if a household produces surplus maize for sale. The final column of Table 7.4 reveals whether the interviewed households brought any maize to market in 2011.

Despite the support provided by ADAM for maize production, it is clear that supported households cultivated a *smaller* area with maize than comparison households and subsequently harvested a smaller quantity. There is no difference in the average maize yield between the supported and comparison households, nor in the proportion (seven per cent) who sold some maize in 2011.

Table 7.4: Production of maize by household members

	Quantity of maize produced in 2011 (quintals)	Area cultivated with maize in 2011 (cuerdas)	Maize yield on 2011 (quintals per cuerda)	Any maize sold during 2011 (binary variable)
<i>Unadjusted:</i>				
Sample mean	6.508	5.576	1.694	0.068
Intervention mean	5.740	4.610	1.752	0.068
Comparison mean	6.989	6.199	1.659	0.068
Unadjusted difference	-1.250 (-1.61)	-1.590*** (-2.87)	0.093 (0.54)	-0.000 (-0.00)
Observations:	379	365	341	382
<i>PSM (ATT)</i>				
Post-matching difference (kernel)	-2.213*** (-2.73)	-1.962*** (-3.16)	0.101 (0.61)	-0.004 (-0.14)
Observations:	368	363	330	371
Post-matching difference (no replacement)	-1.851** (-2.31)	-1.851*** (-2.88)	0.034 (0.18)	-0.021 (-0.79)
Observations:	368	362	330	371
<i>Multivariable Regression:</i>				
MVR/probit coefficient (fixed effects; robust standard errors)	-2.077*** (-2.97)	-2.045*** (-5.91)	0.113 (0.76)	0.002 (0.09)
Observations:	339	343	319	341
MVR coefficient (robust regression)	-2.077*** (-2.97)	-2.045*** (-5.91)	0.113 (0.76)	–
Observations:	339	343	319	
MVR/probit coefficient with control functions (robust SE)	-2.167*** (-3.12)	-2.132*** (-6.14)	0.089 (0.60)	0.001 (0.04)
Observations:	337	341	317	341

t statistics in parentheses

* p<0.1, ** p<0.05, *** p<0.01

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented.

The lower concentration on maize may be a consequence of ADAM's activities supporting cultivation of other crops. However, the lack of an improvement in yield is surprising, since many of the beneficiaries have received fertiliser and technical support specifically intended to make maize production more efficient. The interaction test in Table 7.5 further confirms that there is no evidence for an effect on maize yields among those who received fertiliser, even when distinguished from the supported producers who did not.

Maize yields are no higher among the supported producers, even for those who received fertiliser as part of the project.

Table 7.5: Results of fertiliser distribution interaction test for maize yield (in quintals per cuerda), regressed on intervention dummy variable

Original intervention coefficient	Intervention coefficient with intervention × fertiliser distribution interaction variable	Coefficient on intervention × fertiliser distribution interaction variable
0.136 (0.91)	0.171 (0.90)	-0.076 (-0.27)

t statistics in parentheses

* p < 0.1, ** p < 0.05, *** p < 0.01

Coefficients for covariates not presented.

The coefficient in the first column differs from that presented in Table 7.4, because three observations where the status of having received fertiliser or not is not known.

A major aim of the GUAB49 project was to encourage the production of broccoli for export. Table 7.6 shows results for whether households produced and sold any broccoli during 2011. (Note that the survey did not collect data on the *quantity* of production or sales.) These data confirm that the project was apparently successful in encouraging production of broccoli. The interaction test in Table 7.7 further reveals that the effect of increasing broccoli production is confined to those supported producers who were specifically identified as having participated in GUAB49 (as opposed to having participated in some of the activities of the earlier project, GUAB38).

Table 7.6: Household produced and sold broccoli during 2011

	Produced broccoli	Sold broccoli
<i>Unadjusted:</i>		
Sample mean		
Intervention mean	0.374	0.306
Comparison mean	0.216	0.116
Unadjusted difference	0.159*** (3.40)	0.190*** (4.71)
Observations:	379	379
<i>PSM (ATT)</i>		
Post-matching difference (kernel)	0.142*** (2.64)	0.175*** (3.64)
Observations:	370	370
Post-matching difference (no replacement)	0.140*** (2.66)	0.203*** (4.34)
Observations:	370	370
<i>Multivariable Regression:</i>		
Probit coefficient (fixed effects; robust standard errors)	0.163*** (3.17)	0.175*** (4.00)
Observations:	339	339
Probit coefficient with control functions (robust SE)	0.167*** (3.26)	0.175*** (4.03)
Observations:	339	339

t statistics in parentheses
 $p < 0.05$, $** p < 0.01$, $*** p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

Producers supported under the GUAB49 project were much more likely to grow and sell broccoli during 2011 than the comparison producers.

Table 7.7: Results of GUAB49 interaction test for broccoli production and sales, regressed on intervention dummy variable

	Original intervention coefficient	Intervention coefficient with intervention x broccoli intervention interaction variable	Coefficient on intervention x broccoli intervention interaction variable
Broccoli produced	0.163*** (3.17)	-0.016 (-0.21)	0.350*** (3.46)
Broccoli sold	0.175*** (4.00)	0.031 (0.49)	0.270*** (3.11)

t statistics in parentheses
 $* p < 0.1$, $** p < 0.05$, $*** p < 0.01$
 Coefficients for covariates not presented.

Respondents were also asked about the range of other crops they produced and sold during 2011. Table 7.8 shows the results for the numbers of crops produced and sold after maize and broccoli are excluded. The projects do appear to have had a positive effect on the household's crop diversity, though

Supported producers also produced a wider range of crops at a commercial level, particularly vegetables.

the effect is much clearer on the number of crops brought to market than on the number of crops grown. Examination of the disaggregated data (not shown here) shows that most of the differences are in the production of vegetables – especially carrots, radishes, and *güicoy* (a type of squash) – rather than on more traditional staple crops. This could be seen as consistent with the fact that some supported producers received training in kitchen gardening. However, the third and fourth columns of Table 7.8 show that the effect remains even when those participants in the kitchen garden activity are excluded from the analysis. This effect, then, appears to be due to the wider agricultural extension support which ADAM staff provided to association members through these projects.

Table 7.8: Number of crops (other than maize and broccoli) produced and sold by the household in 2011

	All respondents		Excluding participants in the kitchen-garden activity	
	Number of crops produced	Number of crops sold	Number of crops produced	Number of crops sold
<i>Unadjusted:</i>				
Sample mean	2.768	1.601	2.831	1.711
Intervention mean	2.973	1.959	3.228	2.404
Comparison mean	2.638	1.374	2.638	1.374
Unadjusted difference	0.335 (1.32)	0.585** (2.57)	0.590** (2.15)	1.029*** (4.11)
Observations:	383	383	349	349
<i>PSM (ATT)</i>				
Post-matching difference (kernel)	0.261 (0.97)	0.578** (2.40)	0.460 (1.42)	0.807*** (2.73)
Observations:	372	372	338	338
Post-matching difference (no replacement)	0.574** (2.07)	0.796*** (3.33)	0.692** (2.13)	1.120*** (3.74)
Observations:	372	372	338	338
<i>Multivariable Regression:</i>				
MVR coefficient (fixed effects; robust standard errors)	0.169 (0.64)	0.458** (2.03)	0.240 (0.82)	0.638** (2.34)
Observations:	343	343	310	310
MVR coefficient (robust regression)	0.020 (0.09)	0.195 (1.43)	0.146 (0.60)	0.249 (1.50)
Observations:	343	343	310	310
MVR coefficient with control functions (robust SE)	0.188 (0.72)	0.480** (2.15)	0.244 (0.84)	0.639** (2.34)
Observations:	341	341	308	308

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented.

The kitchen gardening activity was particularly aimed at women, so it is of interest to examine whether it effected women's engagement in agriculture. Table 7.9 reports results for the binary variable which is positive when at least one woman in the household is engaged in agriculture and zero otherwise. As can be seen, there is no difference between the supported and comparison households in this regard.

Table 7.9: Any adult women in the household are engaged in agriculture

<i>Unadjusted:</i>	
Sample mean	0.285
Intervention mean	0.324
Comparison mean	0.260
Unadjusted difference	0.065 (1.37)
Observations:	383
<i>PSM (ATT)</i>	
Post-matching difference (kernel)	0.039 (0.80)
Observations:	377
Post-matching difference (no replacement)	0.048 (0.94)
Observations:	377
<i>Multivariable Regression:</i>	
Probit coefficient (fixed effects; robust standard errors)	0.025 (0.49)
Observations:	341
Probit coefficient with control functions (robust SE)	0.026 (0.50)
Observations:	341

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented.

There are no indications of impact from the kitchen gardening intervention, in terms of diversity of crops produced or women's engagement in agriculture.

7.3.2 Household income and wellbeing

Various measures were collected in the survey that can be used to evaluate household consumption, wealth and poverty. The simplest measure involved asking respondents about their ability to meet basic household needs, without relying on savings, selling assets, or external support. Table 7.10 shows that nearly two-thirds of respondents responded positively to this question overall. The unadjusted figures show that the proportion of positive responses was slightly higher among the supported households than the comparison households. However, once the baseline differences are controlled for, this difference disappears.

Table 7.10: Proportion of respondents reporting that their household is able to meet its basic needs and save for the future from household income

Supported producers are no more likely to report that they can meet their basic household needs than are comparison producers.

	Overall
<i>Unadjusted:</i>	
Sample mean	0.645
Intervention mean	0.689
Comparison mean	0.617
Unadjusted difference	0.072 (1.44)
Observations:	383
<i>PSM (ATT)</i>	
Post-matching difference (kernel)	0.017 (0.34)
Observations:	377
Post-matching difference (no replacement)	0.000 (0.00)
Observations:	377
<i>Multivariable Regression:</i>	
Probit (fixed effects; robust standard errors)	0.070 (1.25)
Observations:	345
Probit with control functions (robust SE)	0.083 (1.44)
Observations:	343

t statistics in parentheses
^{*} $p < 0.05$, ^{**} $p < 0.01$, ^{***} $p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

A more sophisticated indicator of household income is to examine the value of expenditure reported by households. As described in Section 5.2, the survey asked all respondents to estimate the value of all the food items members of their households consumed in the past seven days, as well as all other regular and non-regular expenditures made both in the past month and year, respectively. These expenditure details were aggregated and converted into a per-person per-day figure. Since the consumption figures have a large range with a few extreme figures, they were also transformed onto a natural logarithmic scale, in order to both normalise the distribution and minimise the influence of extreme values. The results after this transformation are shown in Table 7.11.

Table 7.11: Value of household consumption expenditure (natural logarithm of quetzals per person per day)

<i>Unadjusted:</i>	
Sample mean	3.034
Intervention mean	3.078
Comparison mean	3.006
Unadjusted difference	0.072 (1.54)
Observations:	383
<i>PSM (ATT)</i>	
Post-matching difference (kernel)	0.053 (1.12)
Observations:	376
Post-matching difference (no replacement)	0.051 (0.99)
Observations:	376
<i>Multivariable Regression:</i>	
MVR coefficient (fixed effects; robust standard errors)	0.024 (0.57)
Observations:	345
MVR coefficient (robust regression)	0.038 (0.91)
Observations:	345
MVR coefficient with control functions (robust SE)	0.027 (0.65)
Observations:	343

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented.

There are no detectable differences in the level of household expenditure between supported and comparison households.

The estimates of per-capita expenditure are slightly higher among supported households than comparison households, but these differences are not statistically significant. There is, then, little evidence that expenditure – and by implication, household income – has increased as a result of the projects.

Table 7.12 uses the survey questions on food security to derive an alternative measure for impact on household wellbeing. Note that in this table, zero is the food security level of the average respondent and higher figures represent *lower* food security. In general, reported levels of food security are good, with only a minority of respondents reporting that members of their households had not been able to eat sufficiently during the four weeks prior to the survey. Table 7.12 shows that there are no observable differences in terms of food security between the supported and comparison households. However, this analysis is not particularly meaningful, given the lack of variation in the data on this particular measure: most household in both the intervention and comparison groups reported not having any problems related to household food security.

Table 7.12: Food security score (first principle component – higher numbers represent lower food security)

	Overall
<i>Unadjusted:</i>	
Sample mean	0.000
Intervention mean	0.003
Comparison mean	-0.002
Unadjusted difference	0.005 (0.03)
Observations:	376
<i>PSM (ATT)</i>	
Post-matching difference (kernel)	0.016 (0.08)
Observations:	374
Post-matching difference (no replacement)	0.281 (1.46)
Observations:	374
<i>Multivariable Regression:</i>	
MVR coefficient (fixed effects; robust standard errors)	-0.001 (-0.01)
Observations:	338
MVR coefficient (robust regression)	-0.032 (-0.26)
Observations:	338
MVR coefficient with control functions (robust SE)	-0.018 (-0.10)
Observations:	336

Similarly, there is no difference between supported and comparison households in the reported levels of food security.

t statistics in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

The kitchen gardens intervention was intended to increase dietary diversity. The data collected on food consumption allow an evaluation of how diverse each household’s diet is. Table 7.13 shows the number of food types reported to have been consumed in the household during the seven days prior to the survey.

It is clear from Table 7.13 that the supported households reported consuming a wider range of food types than the comparison households. For example, the supported households consumed, on average, five types of fruits and vegetables in the week prior to the survey, compared to four types in the comparison households. If the greater diversity of food types was a result of the kitchen gardens, this should be confined to the consumption of vegetables, but in fact it applies more widely. In particular, the first column of Table 7.13 shows that there is a significant difference overall in food types consumed, and there are corresponding differences when considering only carbohydrates, pulses, meat and fish, or drinks (full data are not shown here). The third and fourth columns of Table 7.13 demonstrate that this finding does not appear to be related to the kitchen-gardens activity; the effect remains (and the estimates of the effect sizes are almost unchanged) when the kitchen garden participants are removed from the analysis.

Table 7.13: Number of food types consumed by household members in the 7 days prior to the survey

	All respondents		Excluding participants in the kitchen-garden activity	
	All food types	Types of fruit and vegetables	All food types	Types of fruit and vegetables
<i>Unadjusted:</i>				
Sample mean	15.809	4.366	15.814	4.375
Intervention mean	17.581	4.919	18.123	5.114
Comparison mean	14.694	4.017	14.694	4.017
Unadjusted difference	2.887*** (5.86)	0.902*** (4.76)	3.429*** (6.36)	1.097*** (5.30)
Observations:	383	383	349	349
<i>PSM (ATT)</i>				
Post-matching difference (kernel)	2.342*** (4.76)	0.649*** (3.42)	2.634*** (4.63)	0.669*** (2.89)
Observations:	374	374	341	341
Post-matching difference (no replacement)	2.157*** (4.24)	0.645*** (3.24)	2.824*** (4.72)	0.642*** (2.64)
Observations:	373	373	341	341
<i>Multivariable Regression:</i>				
MVR coefficient (fixed effects; robust standard errors)	2.172*** (4.54)	0.625*** (3.37)	2.397*** (4.31)	0.648*** (3.01)
Observations:	345	345	312	312
MVR coefficient (robust regression)	2.135*** (4.36)	0.644*** (3.49)	2.595*** (4.63)	0.747*** (3.56)
Observations:	345	345	312	312
MVR coefficient with control functions (robust SE)	2.211*** (4.58)	0.633*** (3.39)	2.385*** (4.24)	0.604*** (2.78)
Observations:	343	343	310	310

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented.

*Supported households apparently have a more diverse diet than comparison households – but this difference is **not** a result of kitchen garden promotion.*

Data were also collected through the survey on various household wealth indicators such as the characteristics of the household's dwelling and asset ownership. This measure is likely to better reflect more established household wealth position, in comparison with current consumption or food security, which tend to reflect current and recent income. As described in Section 5.2, the survey asked about ownership of a series of assets and other wealth indicators, both in 2010 and on the date of the survey, and these observations were used to create an index of the change in the ownership of each asset. The effects on the resulting index of changes in asset ownership are shown in Table 7.14. In this table, higher positive numbers represent greater household wealth compared to the rest of the sample.

Not surprisingly given the previous results, Table 7.14 does not reveal any difference between the supported and comparison households in terms of these wealth indicators.

Table 7.14: Change in asset index between 2010 and date of survey

	Overall
<i>Unadjusted:</i>	
Sample mean	0.002
Intervention mean	-0.088
Comparison mean	0.057
Unadjusted difference	-0.145 (-0.61)
Observations:	378
<i>PSM (ATT)</i>	
Post-matching difference (kernel)	-0.072 (-0.43)
Observations:	375
Post-matching difference (no replacement)	-0.129 (-0.75)
Observations:	372
<i>Multivariable Regression:</i>	
MVR coefficient (fixed effects; robust standard errors)	-0.162 (-0.85)
Observations:	359
MVR coefficient (robust regression)	-0.031 (-0.42)
Observations:	357
MVR coefficient with control functions (robust SE)	-0.121 (-0.64)
Observations:	357

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented

There is no evidence that the assessed interventions increased household wealth status.

7.3.3 Attitudes towards gender roles

The remainder of the outcome measures are derived from the women's questionnaire. As discussed in Section 6.1, the women's questionnaire was administered to the primary respondent when she was female (i.e. those women who were themselves direct beneficiaries and the appropriate comparison respondents) and with the wife or partner of primary respondent when he was male (i.e. in the case of male direct beneficiaries and the appropriate comparison respondents). The table that follows displays the disaggregated results for these two groups, thereby, revealing differences in the effects on women as direct beneficiaries and as indirect beneficiaries of the projects.

First, we will examine the effect on women's attitudes to women's economic roles. As described in Section 5.1, all respondents were asked to state their level of agreement or disagreement with a series of 15 statements on the economic roles of women, and an overall score was constructed using factor analysis. These scores range between 0 and 1, with a higher score representing more positive attitudes to women's roles. Table 7.15 presents the results.

Table 7.15: Attitudes to gender roles among respondents (scores constructed by factor analysis)

	Overall	Female respondents	Wives/partners of male respondents
<i>Unadjusted:</i>			
Sample mean	0.000	0.000	0.000
Intervention mean	0.135	0.089	0.215
Comparison mean	-0.088	-0.056	-0.148
Unadjusted difference	0.223**	0.146	0.364**
	(2.32)	(1.23)	(2.20)
Observations:	368	238	130
<i>PSM (ATT)</i>			
Post-matching difference (kernel)	0.280***	0.145	0.519***
	(2.75)	(1.24)	(2.68)
Observations:	366	238	128
Post-matching difference (no replacement)	0.218**	0.110	0.475**
	(1.98)	(0.85)	(2.27)
Observations:	360	238	122
<i>Multivariable Regression:</i>			
MVR coefficient (fixed effects; robust standard errors)	0.135	0.071	0.353*
	(1.28)	(0.55)	(1.67)
Observations:	327	209	118
MVR coefficient (robust regression)	0.157	0.165	0.400**
	(1.45)	(1.19)	(2.03)
Observations:	327	208	118
MVR coefficient with control functions (robust SE)	0.156	0.101	0.360*
	(1.47)	(0.80)	(1.83)
Observations:	322	206	112

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented.

The wives and partners of male beneficiaries have better attitudes to women's economic roles than do comparison producers.

While the results of the different statistical tests vary, the balance of evidence suggests an overall positive effect of the projects on attitudes towards women's economic roles. This effect is much more pronounced among the wives and partners of male beneficiaries than among the direct female beneficiaries. (The effect among female respondents is not statistically significant.)

It should be noted that the female beneficiaries do express more positive attitudes to women's economic roles than the wives and partners of male beneficiaries do. This is expected, since these are women who have taken the initiative to join a producers' association, while most of the wives and partners of male beneficiaries have not. However, the difference with the comparison group in each case – the best estimate of the change which ADAM and these projects have produced – is higher among the wives and partners of male beneficiaries. In part, there are likely fewer signs of impact among the female direct beneficiaries because they were already at a more positive level.

7.3.4 Women's influence in community-level producer associations

Respondents to the women's questionnaire were also asked the extent to which they disagreed with nine statements about their ability to influence affairs in producer associations in their communities. (Note that these questions did not assume that the respondent was a member of such an association, but non-members would presumably disagree with most of the

statements about their influence.) An overall score was again constructed by using factor analysis.

The results, shown in Table 7.16, closely match the results on attitudes to gender roles: there is no significant evidence of impact among the female beneficiaries, but a positive difference among the wives and partners of male beneficiaries was identified. In this case, it is important to remember that the majority of the female beneficiaries (and the appropriate comparison respondents) were members of female-only associations or groups, and they generally have a better level of participation than those in mixed groups. Where ADAM and/or these projects appear to have made a difference is in promoting opportunities for women’s participation in the mixed cooperatives.

**Table 7.16: Women’s influence in community-level producer associations
(scores constructed by factor analysis)**

	Overall	Female respondents	Wives/partners of male respondents
<i>Unadjusted:</i>			
Sample mean	0.000	0.000	0.000
Intervention mean	0.112	0.040	0.233
Comparison mean	-0.072	-0.024	-0.162
Unadjusted difference	0.184** (2.07)	0.064 (0.58)	0.395*** (2.67)
Observations:	375	241	134
<i>PSM (ATT)</i>			
Post-matching difference (kernel)	0.230** (2.32)	0.103 (0.90)	0.445** (2.51)
Observations:	361	232	129
Post-matching difference (no replacement)	0.222** (2.19)	0.077 (0.63)	0.480*** (2.69)
Observations:	362	232	130
<i>Multivariable Regression:</i>			
MVR coefficient (fixed effects; robust standard errors)	0.138 (1.39)	0.093 (0.75)	0.332* (1.83)
Observations:	331	211	120
MVR coefficient (robust regression)	0.150 (1.47)	0.117 (0.89)	0.377** (2.00)
Observations:	331	210	120
MVR coefficient with control functions (robust SE)	0.134 (1.34)	0.090 (0.72)	0.362** (2.02)
Observations:	326	208	113

t statistics in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

There is evidence that women have a greater ability to influence affairs in the supported associations than the comparison groups and associations.

7.3.5 Women’s involvement in household decision-making

A large component of the women’s questionnaire was devoted to asking women about the degree of their involvement in 25 household decision-making areas. Responses were scored from zero (no involvement) to 4 (sole decision-maker), and a percentage score calculated for the woman’s overall involvement.

On this measure, there is no discernable difference between the position of women in the supported households and comparison households. The women who are direct beneficiaries (and the corresponding comparison

respondents) have higher decision-making scores than do the wives and partners of male beneficiaries – but this is not surprising, since many of those women are household heads.

Table 7.17: Women’s involvement in household decision-making (percentage score)

	Overall	Female respondents	Wives/partners of male respondents
<i>Unadjusted:</i>			
Sample mean	0.612	0.631	0.579
Intervention mean	0.619	0.648	0.571
Comparison mean	0.608	0.620	0.585
Unadjusted difference	0.012 (0.92)	0.028 (1.63)	-0.014 (-1.02)
Observations:	376	241	135
<i>PSM (ATT)</i>			
Post-matching difference (kernel)	-0.017 (-1.18)	-0.013 (-0.66)	-0.024 (-1.61)
Observations:	348	227	121
Post-matching difference (no replacement)	-0.001 (-0.07)	0.008 (0.38)	-0.027* (-1.82)
Observations:	348	227	121
<i>Multivariable Regression:</i>			
MVR coefficient (fixed effects; robust standard errors)	-0.007 (-0.70)	-0.002 (-0.15)	-0.009 (-0.63)
Observations:	332	211	121
MVR coefficient (robust regression)	-0.015 (-1.63)	-0.014 (-1.15)	-0.006 (-0.41)
Observations:	332	209	121
MVR coefficient with control functions (robust SE)	-0.005 (-0.51)	0.001 (0.04)	-0.011 (-0.83)
Observations:	327	208	114

t statistics in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

There is no evidence that the project bolstered women’s household decision-making power.

Oxfam GB global indicator for women’s empowerment is derived by combining the measures used in Tables 7.16 and 7.17. This indicator, in particular is, the proportion of women demonstrating greater ability to influence decisions *both* in the household and in community associations than the “typical” comparison woman. In particular, a household was coded with 1 if the female respondent was both above the median of the comparison group on both measures and 0 if otherwise. Since there is no overall difference between the supported and comparison households in women’s involvement in household decision-making, it is clear that there will also be no difference in terms of the global indicator. This is confirmed in Table 7.18.

Table 7.18: Global indicator for women’s empowerment

	Overall	Female beneficiaries	Wives/partners of male beneficiaries
<i>Unadjusted:</i>			
Sample mean	0.271	0.290	0.237
Intervention mean	.	0.304	0.309
Comparison mean	0.249	0.282	0.188
Unadjusted difference	0.057	0.022	0.122
	(1.22)	(0.37)	(1.64)
Observations:	376	241	135
<i>PSM (ATT)</i>			
Post-matching difference (kernel)	0.003	0.001	0.008
	(0.41)	(0.08)	(0.63)
Observations:	349	234	115
Post-matching difference (no replacement)	0.002	0.004	0.002
	(0.28)	(0.41)	(0.18)
Observations:	353	234	119
<i>Multivariable Regression:</i>			
Probit coefficient (fixed effects; robust standard errors)	-0.086	-0.230	0.430
	(-0.52)	(-1.08)	(1.37)
Observations:	324	207	116
Probit with control functions (robust SE)	-0.049	-0.206	0.503
	(-0.29)	(-0.97)	(1.55)
Observations:	322	205	114

t statistics in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

The findings are not positive for Oxfam’s global indicator for women’s empowerment.

7.3.6 Women’s self-efficacy

The General Self-Efficacy Scale was also administered to the female respondents. This involved asking them about their ability to solve problems and cope with particular situations. An overall score was generated for each respondent with factor analysis, and the differences with respect to the intervention and comparison women are presented in Table 7.19. For this measure, there is strong evidence of a positive effect – and this is for both groups of beneficiaries. The interaction test in Table 7.20 confirms that the effect on self-efficacy is greater among the wives and partners of male beneficiaries. Nevertheless, there is still a significant effect among the direct women beneficiaries themselves.

**Table 7.19: Women’s self-efficacy
(scores constructed by factor analysis)**

	Overall	Female respondents	Wives/partners of male respondents
<i>Unadjusted:</i>			
Sample mean	0.000	0.000	0.000
Intervention mean	0.273	0.199	0.397
Comparison mean	-0.173	-0.122	-0.268
Unadjusted difference	0.445***	0.321***	0.664***
	(4.74)	(2.70)	(4.34)
Observations:	374	240	134
<i>PSM (ATT)</i>			
Post-matching difference (kernel)	0.374***	0.369***	0.385*
	(3.46)	(2.99)	(1.77)
Observations:	358	237	121
Post-matching difference (no replacement)	0.339***	0.256**	0.533***
	(3.24)	(2.07)	(2.83)
Observations:	358	237	121
<i>Multivariable Regression:</i>			
MVR coefficient (fixed effects; robust standard errors)	0.394***	0.222*	0.620***
	(3.70)	(1.77)	(3.63)
Observations:	331	211	120
MVR coefficient (robust regression)	0.405***	0.232*	0.610***
	(3.81)	(1.73)	(3.11)
Observations:	330	210	119
MVR coefficient with control functions (robust SE)	0.370***	0.205	0.647***
	(3.51)	(1.65)	(3.90)
Observations:	326	208	113

t statistics in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
 PSM estimates bootstrapped 1000 repetitions.
 Coefficients for covariates used are not presented.

Women’s self-efficacy is significantly higher among supported producers than comparison producers.

Table 7.20: Results of interaction test for women’s self-efficacy, regressed on female direct beneficiary dummy variable

	Original intervention coefficient	Intervention coefficient with intervention x female direct beneficiary interaction variable	Coefficient on intervention x female direct beneficiary interaction variable
Broccoli sold	0.394*** (3.70)	0.642*** (3.71)	-0.391* (-1.87)

t statistics in parentheses
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
 Coefficients for covariates not presented.

7.3.7 Gender balance in household expenditure

When listing their recent household expenditure (which led to the analysis discussed in Section 7.3.2), respondents were asked, wherever possible, to specify the amounts spent on items for males and females in the household separately. These data were used to examine the ratio between expenditure on males and expenditure on females in the household. The logarithm of the ratios is shown in Table 7.21. The fact that all of the estimates are positive shows that, on average, more was spent on males than females in these households. There is no detectable difference, overall, in this ratio between households supported by the projects and the comparison households. Among the male project beneficiaries, however, there is slight sign of positive impact in this area, though this result is not statistically significant, and so cannot be stated with confidence.

Table 7.21: Ratio of household expenditure on goods and services for males to females (natural logarithm)

	Overall	Female respondents	Wives/partners of male respondents
<i>Unadjusted:</i>			
Sample mean	0.565	0.529	0.619
Intervention mean	0.608	0.497	0.773
Comparison mean	0.537	0.550	0.518
Unadjusted difference	0.071	-0.053	0.256
	(0.54)	(-0.28)	(1.50)
Observations:	301	180	121
<i>PSM (ATT)</i>			
Post-matching difference (kernel)	0.060	-0.038	0.243
	(0.38)	(-0.20)	(0.92)
Observations:	286	178	108
Post-matching difference (no replacement)	0.125	0.007	0.356
	(0.83)	(0.04)	(1.56)
Observations:	286	178	108
<i>Multivariable Regression:</i>			
MVR coefficient (fixed effects; robust standard errors)	0.080	-0.072	0.314
	(0.61)	(-0.37)	(1.61)
Observations:	271	160	111
MVR coefficient (robust regression)	0.162	0.052	0.328*
	(1.32)	(0.28)	(1.84)
Observations:	271	159	110
MVR coefficient with control functions (robust SE)	0.085	-0.075	0.292
	(0.64)	(-0.41)	(1.53)
Observations:	269	159	109

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

PSM estimates bootstrapped 1000 repetitions.

Coefficients for covariates used are not presented.

Household expenditure is generally biased towards males, with no significant differences between the supported and comparison households.

8 Conclusion and Programme Learning Considerations

8.1 Conclusions

The Guatemala Highlands projects have carried out a range of interventions, each involving different subgroups of beneficiaries, with varying degrees of overlap. This effectiveness review finds that the results achieved – at least thus far – are mixed.

The majority of the supported households confirmed that they received various forms of technical and practical support during 2011, and they are also much more likely than the comparison producers to have adopted improved agricultural practices. Unfortunately, this appears to have had little impact to date on their agricultural output. In particular, there is no difference between supported and comparison households in maize yields, even among the 50 per cent of supported producers who were given fertiliser. There is further no evidence that the projects have been successful in encouraging women to cultivate kitchen gardens, though the small number of women who were trained in this area would make any effect difficult to detect. One particularly noteworthy result is that the supported producers cultivated less maize than the comparison producers, but grew a slightly wider range of commercial crops.

There is no significant evidence from any of the various measures of household wellbeing that supported households were better off or had higher income than the comparison households. However, it may be of interest that, while the level of food consumption is approximately the same, supported households reported consuming a greater range of food types at the time of the survey. The reasons for this are not clear from the survey data.

It is important to reinforce that the infrastructure investments made under the GUAB49 project had only recently been completed at the time of the survey. These investments were expected to result in an increase in productive efficiency and hence bolster producer income, but this could not be tested in the effectiveness review.

There is, nevertheless, evidence that the projects have, thus far, worked to improve attitudes towards women's economic roles and created more opportunities for women's participation in community-level producer associations. It is also of interest that evidence for impact is greatest for the wives and partners of the male beneficiaries. However, there is also evidence that the projects have bolstered the self-efficacy of the direct female beneficiaries as well. Unfortunately, there are no indications that the projects have positively affected women's involvement in household level decision-making.

8.2 Programme Learning Considerations

- **Consider whether and how project activities are intended to reinforce each other when designing projects, as well as how evaluation will be carried out.**

The activities of the various projects were intended to be complementary to some extent. However, the groups of beneficiaries who participated in each activity only partially overlapped. For instance, two-thirds of those interviewed as beneficiaries were direct participants in only one of the various activities across the GUAB38 and GUAB49 projects.

To an extent, the two projects, GUAB38 and GUAB49, have acted as a preparatory phase for the GUAB62 project, allowing the programme to experiment with a variety of approaches. However, if experimentation is the aim, it is important to build in means to assess the success of the various activities being implemented. In this case, the structure of the projects made evaluation difficult. This meant that some of the apparent impacts (such as that on crop diversification) cannot be attributed to any of the specific project activities. Some of the activities – notably the women's economic leadership training – were carried out on too small scale to allow for quantitative evaluation of impact.

We recommend that the ways in which projects are expected to generate their intended impacts be more clearly specified, as well as what the relationships between these activities are, and – particularly if the project involves piloting or testing activities – how they will be evaluated.

- **Consider using the sample interviewed for the project effectiveness review as a baseline for the full project (GUAB62).**

Since implementation of the full project (GUAB62) began around the time of the effectiveness review field work, it may be useful to treat the sample of participant and comparison households interviewed as a baseline for this project. In spite of the evidence found for differences in some outcome measures between the supported and comparison households, the two groups are similar enough that they could function as a reasonable baseline. A final evaluation of GUAB62 would then involve re-surveying the same sample of respondents at the end of the full project, and preferably collecting some of the same outcome measures. This would enable a rigorous assessment of this project's impacts to be made.

- **Investigate why the fertiliser distribution and other agricultural support did not lead to improved agricultural productivity or increased household income.**

One of the effectiveness review's clearest findings is that the projects have been successful in encouraging participants to adopt practices such as using improved seeds and disease control methods, and in engaging broccoli production. However, there is no indication that these activities have yet led to an increase in household income. Clearly, the infrastructure investments made under GUAB49 did not have sufficient time to achieve impact at the time of the survey. The difficulties experienced in marketing the broccoli in 2011, despite the alliance with the private companies, may also have prevented that activity from bringing benefits to most households. However, the more general support given to agricultural production over 18 months should normally be expected to result in some modest income gain to beneficiary households. The new marketing conditions agreed more recently with the companies involved in the Alliance may result in improved outcomes in the future.

A particularly surprising finding is that maize yields were no higher among supported producers – even among those who received fertiliser – than the comparison producers. It is not clear from the results presented here why the fertiliser did not lead to increased yields (the survey did not collect data on other agricultural inputs). However, it is important to understand this before making distributions of this kind in the future.