EVALUATION: JUNE 2016 PUBLICATION: OCTOBER 2017

WOMEN'S EMPOWERMENT IN HONDURAS

Impact evaluation of the project Promoting women's engagement in governance in Honduras

Effectiveness Review Series 2016/17

Photo credit:

SIMONE LOMBARDINI and KRISTEN McCOLLUM

OXFAM GB

www.oxfam.org.uk/effectiveness

OXFAM

CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Oxfam in Honduras and local partner teams for being so supportive during this exercise. Particular thanks to: Martiza Gallardo, Sara Martinez, Alvaro Padilla, Suyapa Martinez, Ana Lizeth Cruz, and Mario Alvarado, the consultants who managed the fieldwork.

EXECUTIVE SUMMARY

Oxfam GB's Global Performance Framework is part of the organisation's effort to better understand and communicate its effectiveness, as well as enhance learning across the organisation. Under this Framework, a small number of completed or mature projects are selected at random each year for an evaluation of their impact, known as an Effectiveness Review. The project 'Promoting women's engagement in governance in Honduras' (HONC16) was one of those selected for an Effectiveness Review in the 2016/17 financial year.

Evaluation design

The Effectiveness Review took place in June 2016 in Honduras. It intended to evaluate the success of the 'Promoting women's engagement in government in Honduras' project in achieving its objective of strengthening women's political empowerment. The project was finalized in December 2015, so the evaluation is assessing the impact of the project six months after its conclusion.

The evaluation adopted a mixed-methods approach, employing a quasi-experimental impact evaluation design combined with a qualitative component.

The quantitative impact evaluation aimed at measuring change that is causally attributable to and representative of the project intervention. The evaluation design involved comparing 200 women that had been supported by the project with 292 women in neighbouring communities that had similar characteristics to them in 2008, but who had not participated in the project. A total of 492 women were interviewed using household surveys. At the analysis stage, the statistical tools of propensity-score matching and multivariate regression were used to control for demographic and baseline differences between the households surveyed in project and comparison communities to provide additional confidence when making estimates of the project's impact.

The qualitative component consisted of focus group discussions with project participants in three departments. They served to gather additional information about the project at a national level, and to highlight lessons from aspects of the project that the quantitative portion of the evaluation was unable to address.

Project description

The project began in 2013 and lasted two years. During this period the project directly supported a collection of women's networks in three departments in the west of Honduras (Lempira, La Paz, and Intibuca). It provided space for communication between the regional- and national-level networks, while supporting their organisational capacity. It also promoted dialogue between women's networks and municipal government, lobbied key political actors, and supported efforts to direct municipal funds according to women's needs. Finally, the project provided support to institutionalize and implement strategies aimed at preventing violence against women.

Results

The evaluation found positive and significant results on overall women's empowerment. Quantitative data analysis identified that the project had positive and significant results on various indicators of empowerment at the personal level (such as knowledge of rights and unacceptability of violence), relational-level indicators (such as group decision making and share of household income) and environmental-level indicators (awareness of municipal resources for women). Overall, women in project communities scored positively in 55 per cent of Women's Empowerment Index indicators, compared with 45 per cent for women from comparison communities.

The Effectiveness Review provides evidence that the project had a positive and significant impact on indicators referring to knowledge of individual rights, willingness to defend rights, and willingness to report violence. There is also evidence to suggest that the project increased women's likelihood of deeming violence against her unacceptable, though results in both intervention and comparison groups are very positive. The degree of personal autonomy is high in both the intervention group and the comparison group, though the evaluation finds no significant difference between the two. Finally, there is no evidence of significant impact in terms of other indicators measuring personal empowerment (self-confidence and opinion on women's economic roles), and the low values for each suggest room for improvement.

Regarding relational indicators, participants were significantly more likely than the comparison respondents to have reported influence in group decision making, higher shares of household income, and ability to influence local policies. However, there is no evidence of impact in the other two spheres of decision-making (political and household) examined in the evaluation. Additionally, when asked about their perception of social norms in their community (for example, community members' attitudes towards women's economic or leadership potential) project communities and comparison communities exhibited no difference. However, project participants were more likely to be aware of municipal resources for women, suggesting that the project had some limited positive impact at the environmental level of empowerment.

Summary results of the characteristics of women's empowerment

Level	Characteristic	Linked to project logic	Evidence of impact
Personal	Self-confidence	Yes	No
	Individual knowledge of rights	Yes	Yes
	Unacceptability of violence	Yes	Yes
	Personal autonomy (freedom of movements)	Yes	No
	Willingness to influence and defend her own rights	Yes	Yes
	Willingness to report cases of violence	Yes	Yes
	Opinion on women's economic role	No	No
Relational	Group decision making	Yes	Yes
	Household decision making	Yes	No
	Share household income	No	Yes
	Influencing in local policies	Yes	Yes
	Political decision making	No	No
Environmental	Social Norms	No	No
	Awareness of resources for women	Yes	Yes

The evaluation also investigated additional indicators outside of those used for the Women's Empowerment Index. First, evidence suggests that project participants were more willing to report violence to authorities, but there is no difference between intervention and comparison groups on participants' individual exposure to violence. Project participants were more likely to have heard of violence committed against women close to them, though it is unclear whether this is a result of increased incidence or increased awareness of violence. Project teams should exercise careful monitoring of cases of violence in future projects in order to identify any potential unintended consequences. Regarding political indicators, there is no evidence that political group participation differs between women in the project and comparison

groups, but project participants were more likely to have attended political events such as marches and protests than women from the comparison group.

[Insert info graphic]

Finally, the qualitative component explored the interconnections between national-level influencing and community-level activities. It found some evidence that the national-level activities were instrumental in increasing knowledge of their rights and supported to vocalize their needs and advocate for their rights at local level. Focus group discussions also revealed some up-ward stream of information from local to national, specifically on health, violence, and education. The focus groups also highlighted a need for the national-level activities to be more transparent in their allocation of funds, in order to foster trust and understanding of the use of network resources.

Programme learning considerations

Consider expanding activities promoting women's political influence.

This evaluation finds positive and significant impact on women's ability to influence change, as demonstrated by a range of indicators (for example, knowledge of rights, influence in local policies, and participation in political events) as well as qualitative results. Given the impact of these community-level activities, future projects are encouraged to learn from these successes, so as to encourage women's advocacy in additional community-level policies. These activities could also be broadened in order to achieve policy objectives on a broader geographical level.

Consider revising programme logic to make causal chain for certain goals more explicit.

The evaluation did not provide evidence of positive impact for a few outcomes explicitly stated in the logic model, specifically *self-confidence* and *household decision-making*. However, for contribution to household income – an outcome not as well defined by the theory of change – the project proved very effective. Consider revisiting the logic model, reconsidering assumptions and identifying gaps in the causal chain that may have impeded impact or created unexpected positive change. For example, what additional steps or activities might be necessary to improve self-confidence? What aspects of the project led to an increase in women's contributions to their household income?

Consider altering the targeting strategy.

From the evaluation it emerged that project participants were more likely than the comparison respondents to be already participating in groups and political activities. Additionally, project participants were more highly educated and from wealthier households, on average. The project team is encouraged to explore alternative targeting strategies in order to examine if project activities have even stronger impacts with women who could be considered more vulnerable.

Promote transparency regarding activities and funds used at the national level.

Responses from focus groups highlighted the positive impact of local women's involvement in activities at the national level. Still, some expressed concerns over the use of national-level funds. While this does not necessarily signal misuse, it does suggest that focusing on transparency of funds at the national level and clearly communicating how these finances are allocated could serve to further strengthen the trust and collaboration between networks.

Further explore how women allocated their share of the municipal budget, as these may have caused unexpected positive impacts.

Results suggest a positive impact on women's contribution to household income. While it is still unclear exactly how the project contributed to this outcome, it is likely that project women effectively used the 5 per cent of the municipal budget to fund new economic activities. An increased understanding of the specifics of these activities may allow future projects to replicate this positive impact.

Explore activities to empower women at the household-level

Programme women displayed higher levels of empowerment in *group decision-making*, an outcome explicitly linked to their involvement in women's networks. However, another indicator focusing on women's sphere of influence in the household, *opinion on women's economic role,* showed no effects. Future projects can build on the successes of this programme by targeting household-level power relations to further empower women in various areas of their lives.

1 INTRODUCTION

Oxfam GB's Global Performance Framework is part of the organisation's effort to better understand and communicate its effectiveness, as well as enhance learning across the

organisation. Under this framework, a small number of completed or mature projects are selected at random each year for an evaluation of their impact, known as an Effectiveness Review. One key focus is on the extent to which they have promoted change in relation to relevant OGB global outcome indicators.

The OGB global outcome indicator under which this project has been selected is Women's Empowerment, which is measured by a composite index assessing indicators of empowerment that are relevant to the socio-economic context of the project under analysis. The index is explained in more detail in Section 5.

This Effectiveness Review took place in June 2016 in the departments of Lempira, Intibuca and La Paz in Honduras, and was intended to evaluate the success of the project 'Promoting women's engagement in governance in Honduras' in supporting women to achieve greater empowerment.

The project was implemented by Oxfam's partners: Centro de Estudios de la Mujer (CEM-H), Calidad de Vida and La Asociación de Organismos No Gubernamentales (ASONOG). Building from previous activities conducted implemented since 2008, the project started in 2013 with two levels of intervention. The first, at community level, was directly supporting more than 1,600 women enrolled in women's groups in three departments in the west of Honduras (Lempira, La Paz, and Intibuca). The second, at national level, consisted in conducting advocacy and campaigns interventions.

The key questions for this evaluation were:

- What has been the impact of the project activities implemented at community level in promoting women's empowerment?
- What was the impact of the project in changing other outcome indicators connected with the project logic (such as violence against women, political influence and political participation) among project participants?
- What is the link between the advocacy and campaigns activities conducted by the project at national level and the activities conducted at local level?

It was decided not to assess the impact of the advocacy and campaigns interventions conducted at national level.

Figure 1.1: Map of Honduras

ENTER MAP of Honduras with shaded departments: Lempira, La Paz, Intibuca.

This report presents the findings of the Effectiveness Review, and it is structured as follow. Section 2 briefly reviews the activities and the intervention logic of the project. Section 3 describes the evaluation design used, and Section 4 explains how this design was implemented. Section 5 describes how women's empowerment is measured for this evaluation. Section 6 presents the results of the data analysis, including the descriptive statistics of the population surveyed and the differences in outcome measures between the intervention and comparison groups. Section 7 describes the interconnections between national advocacy and campaigns activities and community level activities. Section 8 concludes with a summary of the findings and some considerations for future learning.

2 PROJECT DESCRIPTION

The overall objective of the 'Promoting women's engagement in governance in Honduras' project was to focus the efforts of various women's networks from four different municipalities, strengthening their capacity to participate in public spaces in order to advocate for their rights and promote a life free of gender-based violence. The project was implemented by three partner organisations: CEM-H, Calidad the Vida and ASONOG.

Contributing to the overall objective, the project intended to:

- Strengthen the organizational capacity of the networks, supporting their engagement in public spaces such as community groups and local government in order to advance their agenda
- Create and strengthen the networks' political dialogue with influential actors, in order to promote decisions that advance women's rights at the local, departmental, regional, and national level
- Ensure that participants in the women's networks have access to mechanisms of attention to, and prevention of, gender-based violence.

Though it represents a continuation of a separate project begun in 2008, the project under evaluation officially started in 2013 and concluded in 2015. During this time, the following activities were implemented:

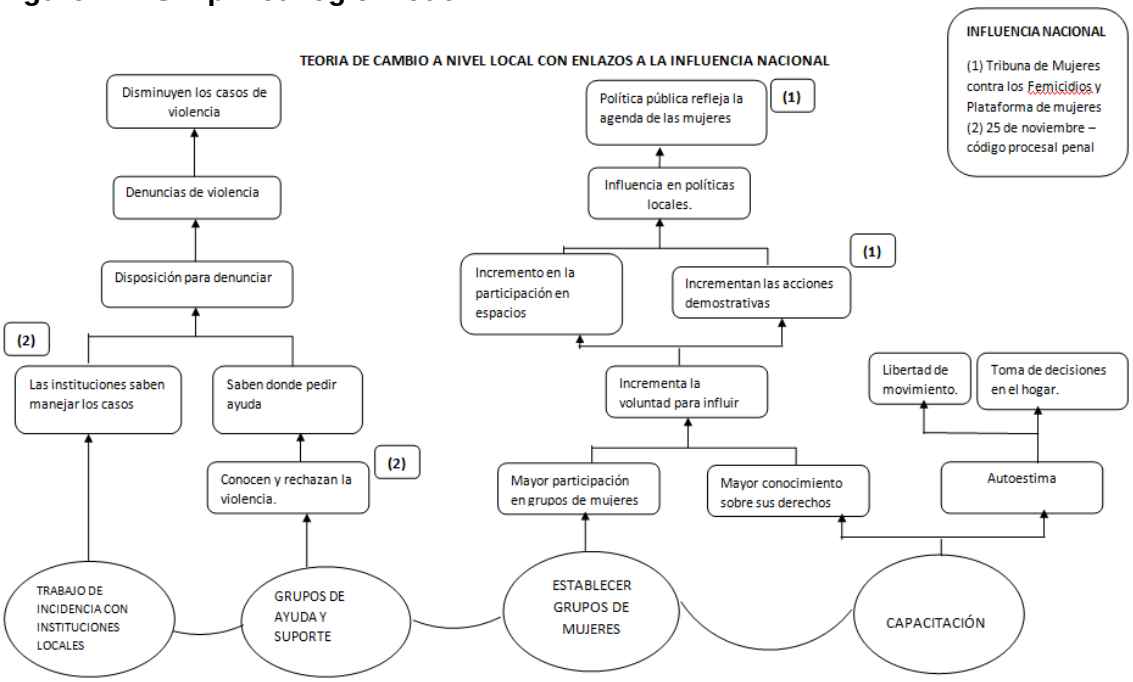
- Support for the creation of a *Presupuesto Municipal Participativo*, a proposal for directing municipal funds according to the wishes of community members
- Lobbying of key political actors
- Hosting of 'dialogue days' between the women's network and local and regional government
- Creation and strengthening of spaces for dialogue and integration of women's networks at the regional and national level
- Establishment of a regional forum with departmental and national authorities to present and agree on networks' strategic demands
- Hosting days for women's networks to institutionalize and implement strategies aiming to prevent gender-based violence.

Figure 2.1 provides a simplified visual representation of the project logic. As depicted, the creation and strengthening women's networks was expected to increase women's participation in local politics and thus result in local-level public policy reflecting the networks' agenda. Additionally, women's increased political involvement was expected to reach a national level, lobbying for federal legislation aligned with the networks' aims of promoting women's rights and preventing gender-based violence.

Additionally, community-level support groups aimed to inform women of how and where to seek support in case of violence, and to encourage reporting such cases. This in turn was expected to decrease the incidence of gender-based violence.

Finally, by providing opportunities for training within the network, the project intended to increase self-esteem, household decision-making power, and freedom of movement for the women participants.

Figure 2.1: Simplified logic model



3 EVALUATION DESIGN

This Effectiveness Review employed two complementary approaches for assessing the impact of the project and answering the evaluation questions. It combined a quantitative quasi-experimental design, which provided results that are representative and can be generalized to the population under analysis, with focus group discussions that provided a deeper understanding of the project's context and interconnections with the national level activities. This section introduces the two approaches.

3.1 QUASI-EXPERIMENTAL DESIGN

The central problem presented in designing an impact evaluation of any social programme is how to compare the outcomes that result from that programme with *what would have been the case* without that programme having been carried out. In the case of this Effectiveness Review, the situation of women in the women's group where the project was implemented was examined through an individual questionnaire – but clearly it was not possible to observe what their situation would have been had they not had the opportunity to participate in this project. In any evaluation, that 'counterfactual' situation cannot be directly observed, it can only be estimated.

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

An ideal approach to an evaluation such as this is to select the units in which the programme will be implemented at random. Random selection minimizes the probability of there being systematic differences between the programme and non-programme units, and so maximizes the confidence that any differences in outcome are due to the effects of the programme.

The project examined in this evaluation conducted its activities at national level as well as local level. The unit at which the programme was implemented at local level was the women's group. Within each of the selected municipalities, women's groups were selected or were specifically established in order for the project activities to be implemented. The selection of municipalities in which to establish these women's groups was not made at random. Instead, project's activities were conducted in municipalities where the partner organizations had previous activities. However, discussions with the implementation staff revealed that there were in fact more municipalities that were considered suitable for project implementation than could be covered by the programme. This allowed a 'quasi-experimental' evaluation design to be adopted, in which the situation of women in those non-implementation municipalities was assumed to provide a reasonable counterfactual for the situation of women in the project groups.

Women in the project groups were 'matched' with women with similar characteristics in non-project (or 'comparison') municipalities. In order to try to control for unobservable variables, such as willingness and motivation to participate in these activities, comparison women were selected from the list of women involved with groups supported by the local Women's Municipal Offices (OMM in Spanish). OMM offices are present in most villages, and women involved in activities implemented by these offices are likely to be women who would have participated in the project, had they had the opportunity.

Matching was performed on the basis of a variety of characteristics – including household size, education level of the woman, whether the head of the household is a woman, group and political participation, productive activities in which the household is engaged, and indicators of material wealth, such as housing conditions and ownership of assets. Some of these characteristics may have been affected by the project itself, so matching was performed on the basis of these indicators *before* the implementation of the project. Since baseline data were not available, survey respondents were asked to recall some basic information about their household's situation from 2008, before the project was implemented. Although this recall data is unlikely to be completely accurate, it should not have led to significant bias in the estimates as long as measurement errors due to the recall data were not significantly different for respondents in the intervention and comparison groups.

The survey data provided a large number of individual and household characteristics on which matching could be carried out. Matching was based on a 'propensity score', which represented the conditional probability of the household being in an intervention village, given particular background variables or observable characteristics. Women in the project and comparison villages were matched based on their having propensity scores within certain ranges. Tests were carried out after matching to assess whether the distributions of each characteristic were similar between the two groups. Details on the validity of the propensity score matching procedure are reported in Appendix 2.

As a check on the validity of the results derived from the propensity-score matching procedure, results were also estimated using alternative PSM models and multivariate regression models. Like propensity-score matching, multivariate regression also controls for measured differences between intervention and comparison groups, but it does so by isolating the variation in the outcome variable explained by being in the intervention group after the effects of other explanatory variables have been accounted for. Appendix 3 provides estimates obtained from these robustness checks.

It should be noted that both propensity-score matching and multivariate regression rely on the assumption that the 'observed' characteristics (those that are collected in the survey and controlled for in the analysis) capture all of the relevant differences between the two groups. If there are 'unobserved' differences between the groups, then estimates of outcomes derived from them may be biased. Unobserved differences between the groups could potentially include differences in attitudes or motivation, differences in community leadership, or other contextual conditions faced by households. The choice of comparison villages, and within each village which women to survey, was made principally to minimize the potential for any such unobservable differences to bias the results. However, results should be interpreted with this assumption in mind, as unobserved differences could still be present between groups.

3.2 QUALITATIVE COMPONENT

The quasi-experimental component of the Effectiveness Review focused on the impact of the community level activities rather than on the national-level influencing component. The country staff expressed the interest in exploring the interconnections between national-level influencing and community-level activities. For this reason this evaluation included three focus group discussions – one in each of the three departments where the project was implemented (Lempira, Intibuca, La Paz).

The objectives for the focus group discussions were to:

1) Explore whether the influencing work conducted at national level strengthened the work conducted at local level
2) Explore whether the work that was developed at the local level was reflected at the national level, and define to what extent the advocacy work carried out by the project at local level was reflected at the national level?
3) Explore the advantages, disadvantages and challenges of developing projects that work simultaneously at local and national level.

The findings from the focus groups on these questions are presented in Section 7.

4 DATA

4.1 SAMPLING OF INTERVENTION AND COMPARISON GROUPS

In the intervention group, from the 66 groups that the project supported in the five municipalities and three departments, 34 groups were randomly selected (with probability proportional to the size of the group). Within those 34 selected groups, nine women were randomly selected from a list of all the women participating in the groups.

Programme staff, partner organizations and the survey team held discussions in order to locate comparable municipalities that were not covered by the project within the same departments. These comparable municipalities, identified as potential comparisons, needed to be located within the same department as the project municipalities and to share similar socio-economic characteristics with them. Seven municipalities were chosen to serve as the comparison group, based on their having

similar socio-economic characteristics, climate, and agricultural production. During fieldwork, in response to cases of violence this was then reduced to five municipalities.

In order to control for endogeneity due to self-selection, respondents within each municipality were randomly selected from lists of women provided by the OMMs. These women were not supported by the project under analysis, though it is likely that they would have participated in the project if it had also been implemented in these municipalities. In cases where a woman was not able available or willing to participate into the survey, the enumerators were instructed to proceed to the following women on the list.

The final sample included 200 project participants (the intervention group) and 292 women selected from OMM lists in comparison municipalities (the comparison group).

4.2 ANALYSIS OF BASELINE CHARACTERISTICS

In order to control for the validity of the comparison group, women in project and comparison communities were compared in terms of their socio-economic characteristics in 2008. These data were based on information recalled during the questionnaire or reconstructed from the household composition at the time of the survey.

To construct a wealth index, we investigate asset ownership as a measure demonstrating more long term changes in wealth for the household. For this reason, respondents were asked about their ownership of various types of household goods and assets, as well as about the condition of their housing. A total of 14 assets and other wealth indicators were used to construct the household wealth index. The wealth indices were then created through applying Principal Component Analysis (PCA) to produce one index of overall wealth. In particular, our wealth index is taken directly from the first principal component. It is then standardized to have a normal distribution.

The full comparison is shown in Table 4.1. It appears that on average women involved in the project are older, the households they are living in were wealthier in 2008, and have more family members. They are also more likely to be the household head, and more likely to participate in groups or political activities.

These differences, which existed before the project, have the potential to bias any comparison of outcomes between the project and comparison villages. It was therefore important to control for these baseline differences when making such comparisons. As described in Section 3, the main approach used in this Effectiveness Review was propensity-score matching (PSM). The full details of the matching procedure applied are described in Appendix 2. After matching, women in the project and comparison villages were reasonably well balanced in terms of the recalled baseline data, with few significant differences between them. However, not all of the women interviewed could be matched, and accordingly 27 project participants and 8 comparison women of the 492 women surveyed had to be dropped from the analysis. As most of the observations dropped came from the intervention group, the evaluation may not be representative of the project's impact on all participants. Rather, results can be interpreted as the project's impact on women with characteristics most similar to non-participant women.

Table 4.1: Baseline characteristics before matching

	Intervention mean	Comparison mean	Difference
1[Lempira department]	0.683	0.679	0.004

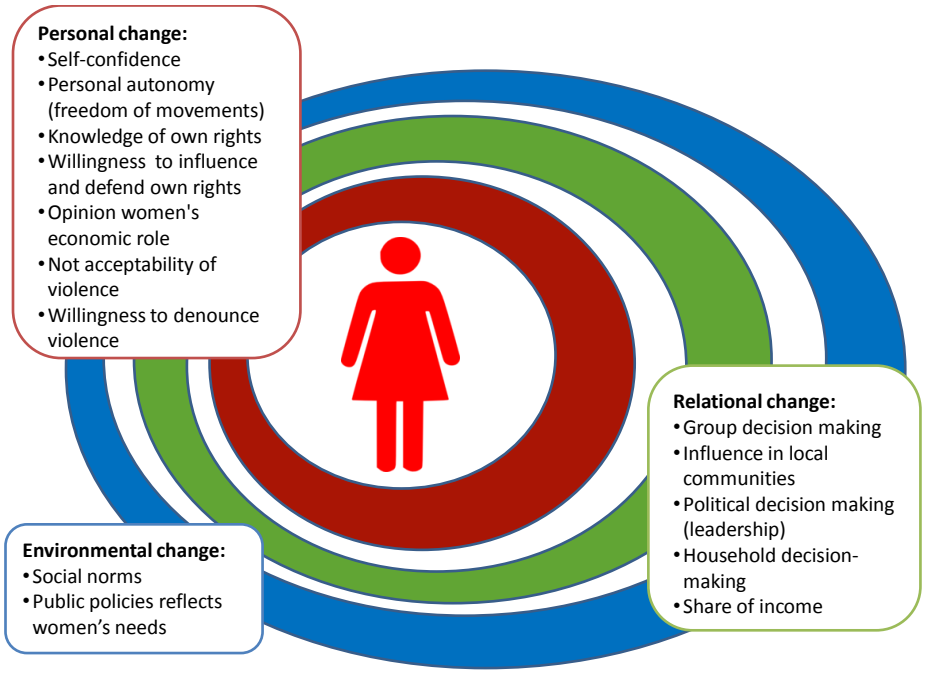
	Intervention mean	Comparison mean	Difference
1[Respondent's ethnic group is Lenca]	0.945	0.990	-0.045***
1[Respondent's religion is Catholic]	0.794	0.587	0.207***
1[Respondent is married]	0.442	0.396	0.046
Respondent's age (years)	40.533	37.874	2.659**
1[Respondent has primary education at least]	0.146	0.099	0.047
1[Respondent is able to work]	0.985	1.000	-0.015**
Number of members of the household	5.583	5.130	0.453**
1[Respondent is household head]	0.342	0.249	0.093**
1[Respondent's contribution to total household income >= 50% in 2008]	0.362	0.222	0.140***
1[Household was in the lowest quintile of the wealth index in 2008]	0.131	0.253	-0.122***
1[Household was in the second quintile of the wealth index in 2008]	0.161	0.222	-0.061*
1[Household was in the third quintile of the wealth index in 2008]	0.216	0.191	0.025
1[Household was in the fourth quintile of the wealth index in 2008]	0.256	0.160	0.096***
1[Respondent worked land or raised animals in 2008]	0.734	0.717	0.017
1[Respondent had formal or temporal employment in 2008]	0.106	0.099	0.007
1[Respondent participated in a community group in 2008]	0.593	0.495	0.098**
1[Respondent participated in a political group in 2008]	0.372	0.201	0.170***

5 MEASURING WOMEN'S EMPOWERMENT IN HONDURAS

The project under review was specifically aimed at promoting women's empowerment as well as increasing household income. In order to to measure women's empowerment Oxfam GB has adopted and adapted an approach that assesses several dimensions of women's empowerment. This approach builds on the Women's Empowerment in Agriculture Index (WEAI) developed by the Oxford Poverty and Human Development Initiative (OPHI) with support from the United States Agency for International Development (USAID) and the International Food Policy Research Institute (IFPRI). Oxfam combined this multidimensional approach with a theoretical framework that defines empowerment as a process where lives are transformed from a situation with limited power to a situation where power is enhanced. Details on the measurement approach can be found in the Oxfam publication "A 'How to' guide to measuring women's empowerment"[1].

This evaluation uses a composite index for measuring women's empowerment, combining indicators associated with changes taking place at three levels: personal, relational, and environmental. Indicators defined at the personal level refer to empowerment taking place within the person. Indicators defined at the relational level refer to empowerment taking place in the power relations within the woman's surrounding network, including within the household and in the community. Finally, indicators defined at the environmental level refer to empowerment taking place in the broader environment, including political space and social norms.

Figure 5.1: Key dimensions of women's empowerment



There is no single generic set of 'women's empowerment' characteristics that are applicable in all contexts. For this reason, the measurement tool employed in each evaluation is designed to be context specific. The choice of indicators used to define women's empowerment was constructed following three steps.

1. During a two-day workshop in Tegucigalpa, programme staff, staff from the partner organisations, a local consultant, a local thematic expert, and the author of the evaluation identified a list of characteristics and indicators aiming to describe an empowered woman in Honduras. A questionnaire was designed and tested, which included questions intended to capture each of the characteristics identified.

2. A group discussion took place with women representing the OMMs from the Lempira department. This step was used to define which of the characteristics and indicators identified at step 1 should be included in the final composite index. Table 5.1 reports the final list of indicators identified in this evaluation and used for constructing the composite index. It is important to note that while not all characteristics considered in this evaluation may be directly linked to the project activities, all were deemed to be important to women's empowerment in the context of western Honduras.

3. In order to combine all the 14 indicators into a composite index, a benchmark was defined for each characteristic in order to identify what it means for a woman to be doing well in relation to the characteristic in question. The composite index measures the proportion of characteristics in which a woman scores positively across the 14 indicators describing empowerment.

We recognise that there is inevitably a degree of arbitrariness in defining such benchmarks. Therefore, Section 6 presents estimates obtained from the same indicators that can be interpreted without applying the binary thresholds, in order to explain in more detail the indicators and levels under analysis. The particular benchmarks used for each characteristic to construct the women's empowerment index are described in Appendix 1.

Table 5.1: Characteristics of women's empowerment examined in this Effectiveness Review

Level	Characteristic
Personal	Self-confidence
	Individual knowledge of rights
	Unacceptability of violence
	Personal autonomy (freedom of movements)
	Willingness to influence and defend her own rights
	Willingness to report cases of violence
	Opinion on women's economic role
Relational	Group decision making
	Household decision making
	Share household income
	Influencing in local policies
	Political decision making
Environmental	Social Norms
	Awareness of resources for women

6 RESULTS

6.1 INTRODUCTION

This report is intended to be free from excessive technical jargon, with more detailed technical information being restricted to the appendices and footnotes. However, there are some statistical concepts that cannot be avoided in discussing the results. In this report, results will usually be stated as the average difference between women living in villages where the project was implemented (the 'intervention group') and the matched women in villages where the project was not implemented (the 'comparison group').

In the tables of results on the following pages, statistical significance will be indicated by asterisks, with three asterisks (***) indicating a *p*-value of less than 1 per cent, two asterisks (**) indicating a *p*-value of less than 5 per cent and one asterisk (*) indicating a *p*-value of less than 10 per cent. The higher the *p*-value, the less confident we are that the measured estimate reflects the true impact. Results with a *p*-value of more than 10 per cent are not considered to be statistically significant.

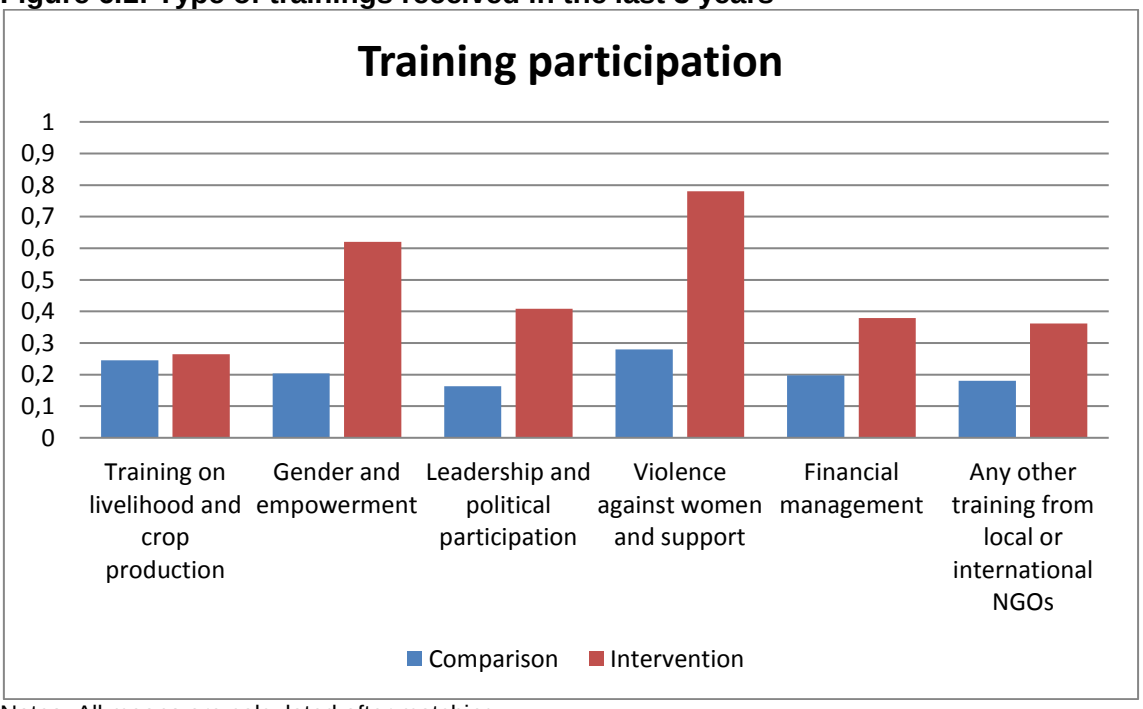
The results are shown after correcting for observable baseline differences between the women interviewed in the intervention group and in the women in comparison villages using a propensity-score matching (PSM) procedure. The details of this procedure are discussed in Appendix 2. All outcomes have also been tested for robustness to alternative statistical models in Appendix 3. Where those alternative models produce markedly different results from those shown in the tables in this section, this is discussed in the text.

6.2 INVOLVEMENT IN PROJECT ACTIVITIES

Before considering the project's effect on outcomes, it is important to examine whether the respondents reported having participated in the activities implemented under this project.

As presented in Section 2 the project supported women in a number of ways. Figure 6.1 shows the proportions of women from intervention and comparison groups who received training on different topics in the three years prior to data collection. Almost 80 per cent of the women interviewed in the intervention group reported having received training on the topic of violence against women and available support, compared with less than 30 per cent in the comparison group. Around 60 per cent of project participants reported having received training on gender and empowerment, compared with less than 20 per cent in the comparison group.

Figure 6.1: Type of trainings received in the last 3 years



Notes: All means are calculated after matching.

Figure 6.1 provides good evidence that the women identified in the comparison group were exposed to a limited extent to activities similar to the one implemented in the intervention group. Since comparison women were selected from women's groups, we should expect to find that these women receive at least a minimal amount of training. However, it is clear that women selected for the intervention group are more heavily involved in activities implemented by the project (covering topics of empowerment, political participation, and support for women experiencing violence). This provides some confidence that we are comparing project participants with women who were exposed only to limited project activities, despite being selected from the OMMs lists.

6.3 IMPACT ON WOMEN'S EMPOWERMENT

This section examines the differences between the intervention and comparison groups on women's empowerment indicators and the overall women's empowerment index described in Section 5.

The outcomes measures examined in this section are:

- Overall women's empowerment.
- Personal level indicators of empowerment.
- Relational level indicators of empowerment.
- Environmental level indicators of empowerment.

6.3.1 Overall women's empowerment

Table 6.1 shows a comparison of the intervention and comparison groups in terms of the measure of women's empowerment presented in Section 5. This measure represents the proportion of characteristics in which women scored positively across the 14 indicators describing empowerment. Estimates suggest that women involved in the project score positively on average in 55 per cent of the indicators, compared with women not involved in the project, who score positively on average on 45 per cent of the indicators. This difference is statistically significant, suggesting that the project has had a positive impact on overall women's empowerment.

Table 6.1: Overall women's empowerment index

	Empowerment index
Intervention group mean:	0.55
Comparison group mean:	0.45
Difference:	0.10***
	(0.04)
Observations intervention:	173
Observations:	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Table 6.2 presents estimates of the index disaggregated under the three levels of empowerment: personal, relational, and environmental. Estimates suggest that there is a positive and significant impact of the project on the dimension of empowerment referring to changes taking place at all three levels.

Table 6.2: Women's empowerment – power dimensions

	Empowerment index (personal)	Empowerment index (relational)	Empowerment index (environmental)
Intervention group mean:	0.60	0.50	0.50
Comparison group mean:	0.50	0.40	0.39
Difference:	0.11***	0.10**	0.12**
	(0.03)	(0.04)	(0.05)
Observations intervention:	173	173	173
Observations:	457	457	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

The following sections present in detail indicators and dimensions included in the index presented above. Figure 6.2 provides a graphical representation of the difference between the intervention and comparison groups for each indicator included in the women's empowerment index.

Figure 6.2: Results on the overall women's empowerment index

[enter info-graphic –ask Chrissy]

6.3.2 Personal

The first level of empowerment refers to changes taking place within the person. This dimension measures changes in personal self-confidence, opinions, attitudes and beliefs. It examines how a woman sees and perceives herself in society, and what capability she has to decide actions and carry them out.

In the context under analysis the following indicators have been identified:

1. Self confidence
2. Personal autonomy (freedom of movements)
3. Knowledge of women's rights
4. Willingness to influence and defend own rights
5. Opinions on women's economic role
6. Not acceptability of gender-based violence
7. Willingness to report violence

Table 6.3: Proportion of women reaching threshold on indicators of personal-level empowerment, by intervention status

	Self-confidence	Personal autonomy	Knowledge of women's rights	Willingness to influence and defend own rights	Opinions on women's economic role
Intervention group:	0.33	0.94	0.51	0.43	0.27
Comparison group:	0.34	0.93	0.16	0.29	0.23
Difference:	-0.02	0.01	0.35***	0.13**	0.03
	(0.07)	(0.03)	(0.04)	(0.06)	(0.08)
Observations intervention:	173	173	173	173	173
Observations:	457	457	457	457	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Self confidence/self-esteem measures the attitude the respondent has towards her-self. Respondents were asked to what extent they agreed or disagreed with the following statements:

- I feel that I'm a person of worth, at least on an equal plane with others.
- I feel that I have a number of good qualities.
- I feel I do not have much to be proud of.
- I am equal to my peers (e.g. sisters, friends, colleagues, etc.).

The first column in Table 6.3 provides estimates on the proportion of women who were considered being self-confident in the intervention and comparison group. On average both intervention and comparison groups present similar levels of self-confidence.

Personal autonomy is the capacity to decide actions for oneself and pursue a course of action in one's life without reference to others. For each of the following decision-making areas, the respondent was first asked who normally makes the decisions about that area .

- Whether you can personally travel to visit relatives outside the community.
- Whether you can personally participate in community group activities or meetings.

A woman scored positively if she identified herself as either solely or jointly responsible for both of these decisions. The second column in Table 6.3 provides evidence that in both groups women appear to have high levels of personal autonomy. There is no significant difference between the two groups.

The third column in Table 6.3 measure whether women have **knowledge of their own rights**. In particular it is testing whether the respondent is aware of the resources from the municipality's budget that should be invested in supporting women, and that she knows the proportion of the total budget. Estimates from the survey suggests that 51 per cent of the women in the intervention group reported the correct answer (5% of the municipality's budget), compared with only 16 per cent in the comparison group. This represents a significant impact of the project in increasing knowledge of women's rights.

Willingness to influence and defend her rights was measured by asking to what extend she agreed or disagreed with the following statements.

- Public meetings held in your village are quite intimidating - it is difficult for a woman like you to stand up and voice any concern.
- If a public decision were made that might negatively affect your life and those of your children, you would not hesitate to stand up and protest despite the possible negative consequences.
- You do not feel intimidated to go and talk to the mayor if some of your rights are not respected.

Column 4 in Table 6.3 shows the proportion of women who disagreed with the first two statements and agreed with the third one. Estimates suggest that there is a difference of 13 percentage points between the intervention and comparison group in this measure.

Opinions of women's economic role measure the respondents' attitudes and beliefs regarding a woman's ability to contribute economically to her household and community. Respondents were asked to what extent they agree or disagree with the following statements.

- Women are just as capable as men of contributing to household income.
- A man's job is to earn money; a women's job is to look after the home and family.

Estimates in the fifth column in Table 6.3 show the proportion of women that strongly agree with the first statement and strongly disagree with the second statement. There appears to be no difference between the intervention and comparison group in this respect.

Table 6.4: Proportion of women reaching threshold on indicators of personal-level empowerment, by intervention status (cont)

	Unacceptability of gender based violence	Willingness to report violence
Intervention group mean:	0.91	0.85
Comparison group mean:	0.80	0.72
Difference:	0.10*	0.12**
	(0.05)	(0.05)
Observations intervention:	173	173
Observations:	457	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Unacceptability of GBV measures the extent to which a woman considers violence against women acceptable. Respondents were asked whether they believed it is acceptable for a man to hit his wife in the following instances:

- If she disobeys her husband or other family members.
- If he suspects that she has been unfaithful.
- If she neglects the children.
- If she spends money without permission.
- If she is not supporting her husband in livestock and agricultural activities.
- If she goes to see her family without the permission of her husband.
- Any other case not mentioned above.

Estimates in the first column in Table 6.4 show the proportion of women that do *not* consider violence acceptable any of these circumstances. On average 91 per cent of women involved in the project do not consider violence acceptable under any of these circumstances, compared with 80 per cent in the comparison group. This difference is statistically significant only at the 10 percent level, meaning that we cannot conclude with confidence whether such a difference is likely to be found among the membership of the project and comparison groups as a whole, or whether it is specific to the particular sample of respondents interviewed[2]. Nevertheless, this provides some tentative evidence that the project had a positive impact on changing attitudes towards violence.

Willingness to report violence measures whether women said that they were willing to report a man who beat them to the authorities. A woman scored positively here if she answered 'yes', and negatively if she answered 'no' or declined to answer. Estimates in Table 6.4 show that 85 per cent of the intervention group reported willingness to report violence committed against them, while only 72 per cent of the comparison group said the same. The difference of 12 percentage points is statistically significant, and suggests that the project increased women's self-declared willingness to report cases of violence against women. However, it is important to note that this does not necessarily measure differences in actual action taken to report violence. We explore this aspect separately in section 6.4.

6.3.3 Relational

This second level of indicators measures empowerment taking place in power relations within the woman's surrounding network. This can be expressed as changes in power relations among individuals. In the context of the project this evaluation identified the following indicators:

1. Household decision-making
2. Share of household income

It also includes indicators that reflect the recognition that empowerment is a collective process, which requires the support and interaction of peers and organisations. The following indicators were identified:

3. Group decision making
4. Influencing in local communities
5. Political decision making (leadership)

Table 6.6 provides estimates for the remaining two indicators of empowerment at the relational level.

Table 6.6: Indicators of relational-level empowerment (power over), by intervention type

	Household decision making [proportion of activities]	Share of household income [proportion of resources]
Intervention group mean:	0.66	0.63
Comparison group mean:	0.63	0.48
Difference:	0.03	0.15*
	(0.07)	(0.08)
Observations intervention:	173	173
Observations:	457	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Involvement in household decision making measures the extent to which the respondent is involved in decisions taking place in her own household. Each respondent was asked who normally makes decisions in the following areas:

- How cash earned from agricultural production is spent.
- Purchase and sale of large animals.
- Purchase and sale of land.
- Purchase of furniture and electronic devices.
- Whether the household should take out a small loan, from what source, how much to borrow, and how to spend the money.
- Whether to migrate, to what place, and when.
- What other non-agricultural business activities the household will engage in, and how to manage these activities.
- How cash earned from these non-agricultural activities is spent.
- What to give relatives when they marry or have a celebration.
- The education of their children.
- Transfer of property to a relative or any other person.

If the woman reported that she is not solely responsible for the decision, then she was asked to what extent she thought she could influence the decision, on a scale from 'not at all' to 'a large extent'. The first column of Table 6.6 reports the proportion of activities in which the respondent makes the decision by herself or in which she is able to influence those decisions to a large extent. In both the intervention and comparison groups, women appear to have high levels of household decision-making power (on average 66 per cent of respondents reported being able to being involved in all household decisions), and there is no significant difference between the two groups.

Share of household income attempts to measure the proportion of income women earn independently from their spouse or other household members.

Respondents were asked:

- Given the total of resources your household consumes (including crop, cash, and services), what was the percentage of your own contribution?

Estimates in the second column in Table 6.6 provide the average proportion of household resources that women in the intervention and comparison groups reported to contribute to the household. On average women in the intervention group contribute 63 per cent of total household income, compared with 47 per cent in the comparison group. The difference is statistically significant different from zero. This result comes as a surprise as the project was not aiming to directly change this outcome indicator.

Conversations with country teams have highlighted one explanation: it is possible that the increased use of the 5 per cent of the municipal budget had an impact on women's income. In some cases, this budget was directed toward economic project fund from which women were able to make and sell handicrafts. Future projects should explore this causal chain as one potential route to achieve women's economic empowerment.

Table 6.5 provides estimates for the second three indicators identified under empowerment at the relational level.

Table 6.5: Indicators of relational-level empowerment (power with), by intervention status

	Group decision making	Influence in local policies	Political decision making
Intervention group mean:	0.70	0.26	0.24
Comparison group mean:	0.51	0.12	0.26
Difference:	0.20***	0.13***	-0.03
	(0.07)	(0.05)	(0.07)
Observations intervention:	173	173	173
Observations:	457	457	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Group decision making measures whether the woman respondent feels moderately or very able to influence group decisions within the women's network, cooperatives, credit groups, farmers' groups or unions, religious groups, and other community groups. Results show a statistically significant difference between intervention and comparison groups, with 70 per cent of project participants expressing abilities to make group decisions, compared to only 51 per cent of women in the comparison group.

Influence in local policies asks whether the respondent has attempted to influence the use of the municipal budget for their communities. Twenty-six per cent of project participants affirmed that they had, while only 12 per cent of comparison respondents expressed having tried to influence local policy in this way. The difference of 13 percentage points is statistically significant.

Finally, **political decision making** measures whether a woman feels moderately or very involved in making important decisions in at least one political body within her community. Results for this indicator do not provide evidence of impact, as women from both intervention and comparison communities report having influenced political decisions at approximately the same rate. It is important to note that this question asked specifically about a range of groups other than women's networks, in order to measure whether their influence has spilled over into other political environments. As such, it does not necessarily measure women's influence within the networks.

6.3.4 Environmental

The final level of empowerment measures aspects of the broader environment. The following indicators were identified for this evaluation:

1. Social norms
2. Awareness of municipal resources for women

Table 6.7 provides estimates of the indicators identified under this dimension.

Table 6.7: Indicators of environmental-level empowerment, by intervention status

	Social Norms	Awareness municipal resources for women
Intervention group mean:	0.27	0.73
Comparison group mean:	0.23	0.54

Difference:	0.05	0.19***
	(0.04)	(0.07)
Observations intervention:	173	173
Observations:	457	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

The **social norms** indicator measures whether respondents feel that men and women from the community have positive attitudes regarding social norms. Respondents were asked how they believed women and men from their communities would respond to the following statements:

- Women are as capable as men at contributing to household income
- Women can be as good leaders as much as men can
- A man's work is to provide income, and a woman's work is to take care of the family and the house

A respondent scored higher in this indicator if she believed community members would respond positively to the statements above (agreeing with the first two statements and disagreeing with the third). Results do not show a difference between intervention and comparison groups on this indicator.

Awareness of municipal resources for women tests whether a woman has heard of resources from the municipal government which are intended to support women. A respondent scored positively in this category if she was aware of these resources. In this case, women from project communities were 19 percentage points more likely to be aware of municipal resources supporting women, and this difference is statistically significant.

6.4 IMPACT ON OTHER OUTCOMES

This section will examine the differences between intervention and comparison groups on other outcome measures relating to the project's logic as discussed in Section 2. The indicators in this section intended to improve conditions for both men and women in the household.

The outcomes measures examined in this section are:

1. Violence against women
2. Political influence

Table 6.4.1 begins the section by reporting on a woman's exposure to violence, reported in two ways:

- Violence experienced by the respondent herself
- Violence experienced by women close to the respondent

6.4.1 Violence against women

Table 6.8: Exposure to violence in the past 12 months, by intervention status

	Exposure to violence (respondent)	Exposure to violence (women close to respondent)
Intervention group mean:	0.28	0.51
Comparison group mean:	0.24	0.37

Difference:	0.05	0.15
	(0.08)	(0.10)
Observations intervention:	150	150
Observations:	383	383[1]

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Results of **exposure to violence (respondent)** in Table 6.8 do not provide evidence of a significant impact of the project. For this indicator, women were asked whether any form of psychological, physical, or sexual violence had been committed against them in the past 12 months. 28 per cent of project participants indicated they had experienced such violence, which does not significantly differ from the reported rate from women in the comparison group.

The indicator of **exposure to violence for women close to the respondent** does not provide clear evidence on whether there is a difference as a result of the project. Women from the intervention group report higher rates of exposure to violence (52 per cent), than their comparison peers (38 per cent). While this difference is not statistically significant, estimates from the robustness checks in Appendix 3 suggest otherwise. Since women from the intervention group did not report significantly higher rates of violence being carried out against them individually, it is likely that if such difference exists, it is due to increased *awareness* of violence against women close to them, rather than an increase in actual incidence of violence. Careful monitoring of cases of violence should be exercised in future projects in order to identify any potential unintended consequences.

Table 6.9: Types of violence experienced in the past 12 months

	Type of violence: psychological	Type of violence: physical	Type of violence: sexual
Intervention group mean:	0.43	0.28	0.13
Comparison group mean:	0.29	0.23	0.08
Difference:	0.15*	0.05	0.05
	(0.09)	(0.07)	(0.03)
Observations intervention:	173	1502	173
Observations:	457	382	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Table 6.9 further elaborates on exposure to violence. Above we present the three types of violence explored (psychological, physical, and sexual) and provide the proportions of women who reported each one. Of the women who reported that either they or women close to them had experienced violence, psychological violence was the most prevalent type in both intervention and comparison groups, followed by physical and then sexual violence. In all types, we see more women from project groups reporting exposure to violence than those from comparison groups. However, these differences are generally not statistically significant. The exception is that for psychological violence, where women from project communities were 15 percentage points more likely to have reported experiencing psychological violence than those

Figures for both indicators were calculated among respondents whose interviews were conducted privately and who were willing to answer questions regarding violence against women. Additionally, the proportion of women who responded was higher in the intervention group than in the comparison group, suggesting there may be some selection bias in the sub-sample of observations for these indicators.

[2] Again, the figure for this indicator was calculated among respondents whose interviews were conducted privately and who were willing to answer questions regarding violence against women. The proportion of women who responded was higher in the intervention group than in the comparison group, suggesting there may be some selection bias in the sub-sample of observations for this indicator.

from comparison communities. Although this difference appears large, it is significant only at the 10 percent level - although consistent with other estimates from Appendix 3. These figures elaborate on the findings from Table 6.8. Though experience of violence did not show a statistically significant difference, it is possible that the rates were high as a result of women being more aware of, and willing to acknowledge psychological abuse as a type of violence.

Table 6.10 examines reported cases of violence to the authorities in the last 3 years.

Table 6.10: Reported violence to an authority in the last 3 years

	Reported violence
Intervention group mean:	0.35
Comparison group mean:	0.22
Difference:	0.12*
	(0.06)
Observations intervention:	172
Observations:	456

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

To measure **reported violence to an authority**, the woman was asked if she knew of any cases in the last 3 years in which a woman reported an act of violence to an authority or support system, such as the police or an OMM. While 35 per cent of participant women had known of reports, only 22 per cent of comparison women had. This difference is also statistically significant, but must be interpreted carefully. As seen previously in Table 6.8, participant women also stated in the survey that they were aware of higher incidence of violence. Considering this, it is difficult to determine whether more frequent reporting is due to higher need of reporting or an increased willingness to report. However, this result does mirror the results of Table 6.4, where women from project communities expressed being more *willing* to report violence than their comparison counterparts.

In any case, Table 6.11 offers some useful summary statistics of the results of project participants' reports of violence.

Table 6.11: Results of project participants' reports of violence

	Number of observations	Average	Std. Dev.
Number of days before response	87	12.18	41.94
Response to report was considered satisfactory	88	0.64	0.48
Action was taken	88	0.59	0.49

Of those who knew of a reported incident of violence, the survey asked for the average time it took to get a response, whether or not the respondent was satisfied with the support received following the report, and whether any action was taken as a result of filing a report. On average, reports of violence received a response in approximately 12 days, although the standard deviation of this figure is large. Response times ranged from only one day to one entire year. When asked about the quality of the support they received following filing a report, 64 per cent of respondents rated it as "good or very good." Finally, 59 per cent of respondents affirmed that action was taken as a consequence of the report that was made.

Table 6.12: Where respondents reported violence, by project status

To whom did you report the violence?	Intervention	Comparison
Legal promoter	.028	.000
Police	.111	.383
Judge	.347	.283
Women's Municipal Offices	.403	.250
Public Prosecutor	.083	.050
Other	.028	.033
Total observations:		
	72	60

Of respondents who knew of reported violence, the authority chosen for reporting differs by intervention type. As seen in table 6.12, project participants and women close to them reported most frequently to the women's municipal office (the location for approximately 40 per cent of reports), while comparison women and those close to them were most likely to report to the police (receiving approximately 38 per cent of reports). These trends are likely related to the difference in willingness to report violence, as the project worked heavily with women's municipal offices.

6.4.2 Political influence

Lastly, we report here on the impact of project activities on participants' political and group influence. Table 6.13 begins with outcomes on political influence and participation.

Table 6.13: Indicators of political influence and participation, by project status

	Started political participation	Event participation	Number of events participated in
Intervention group mean:	0.24	0.84	7.18
Comparison group mean:	0.18	0.67	3.37
Difference:	0.05	0.17**	3.90**
	(0.06)	(0.09)	(1.52)
Observations intervention:	173	173	173
Observations:	457	457	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Started political participation asks whether the respondent is a member of more local political organizations than she was in 2008. This was the case for 21.5 per cent of respondents. There does not appear to be a significant difference in the take-up of political participation between intervention and comparison groups.

A respondent scored positively in **event participation** if she had participated in any political event in the past 3 years. Here we see a statistically significant difference between intervention and comparison groups, with 84 per cent of project participants having participated in such an event, compared to 67 per cent of their comparison peers. Similarly, there appears to be a significant impact of the project on the **number of events participated in**. Women from project communities participated in an average of 7.2 events in the past three years, compared to an average of 3.3 events among women in the comparison group.

Table 6.14: Indicators of group influence and participation, by project status

	Started group participation
Intervention group mean:	0.42
Comparison group mean:	0.36
Difference:	0.06
	(0.06)
Observations intervention:	173
Observations:	457

Notes: Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. All means are calculated after matching.

Table 6.14 examines respondents' participation in community groups, in particular it examines if a woman is involved in more groups now than she was in 2008. Take-up rates do not appear to differ significantly between intervention and comparison groups: 42 per cent of women in project groups were more involved in groups at the time of the survey than in 2008, and 34 per cent of comparison women said the same. However, it is worth noticing that **group decision making** measured in Table 6.5 suggested that women involved in the project have greater group decision-making power, compared to comparison women.

7 INTERCONNECTIONS BETWEEN NATIONAL ADVOCACY AND COMMUNITY ACTIVITIES

In order to explore the interconnections between national level influencing and community level activities, the evaluation also included a qualitative component carried out through three focus group discussions with project participants, one in each of the three departments where the project was implemented (Lempira, Intibuca, La Paz). The qualitative analysis is presented here, organized around three central questions.

Participants in the focus group in Intibuca included 9 members of the executive board of the women's network and 5 interns involved with the network. The La Paz focus group included 12 members of the women's network accompanied by 3 representatives from partner organisations. In Lempira, 10 women from the executive board participated in the focus group. Each discussion was facilitated by a female member of ASONOG.

1) Did influencing work conducted at national level strengthen the work conducted at local level?

In each of the three regions, women responded that national-level organisations were integral in increasing knowledge of their rights and encouraging participation. Women mentioned two main categories of national-level activities: trainings and political activism, each providing a mechanism for local-level empowerment.

For example, women from Lempira mentioned that the national training hosted by ASONOG allowed them to continue empowering themselves through the knowledge of their rights. Women from Intibuca agreed, and also cited the financial support the project provided women for lodging, travel, and meals during their trip as an important national-level contribution. This funding not only allowed them to participate in the training in the capital, but also had the effect of continuing that participation in various community-level groups such as water committees, health volunteering groups, and other community boards. Importantly, participants in the national trainings would replicate those workshops in their own communities. Participants in Intibucá stressed the importance of these replications for spreading awareness on the topics of violence and self-esteem to women who were not able to travel to the capital. Additionally, women from La Paz frequently mentioned the value of the capital visits as a mechanism of knowledge-exchange between other networks:

> *"We have shared what we have been able to achieve and what we haven't, and we see that some [networks] have achieved more than others and they work in one way while other [networks] work differently. But with an exchange, we bring the experiences of how we can do things and they take those [experiences] so they can as well."*

Additionally, all three departments mentioned advocating for the continuation of the law designating 5% of the municipal budget towards women's groups and issues. When the women's networks felt that this law was at risk of being revoked, some representatives of the network travelled to the capital to speak with key influencers about their experiences and the importance of funding. For example, respondents from Intibucá mentioned meeting with representatives from the Gender Commission of the National Congress, the First Lady, and the Director of AHMON (Asociación Hondureña de Municipios). In the latter instance, women from the network urged the Director to ensure that federal funding arrived on time to the municipalities, because they felt that the 5% owed to women would be at risk of being delayed or cut if municipalities were budget-constrained. While it is not possible to attribute a causal relationship between this activism and the subsequent availability of the 5% budget allocation, it is certain that the women *feel* their activism was both effective and crucial. Women from Lempira identify a direct relationship between national-level influencing and local issues, saying:

> "*Speaking about the local-level: sometimes when the transfers [to the municipal budget] are late, it ends up affecting our work as women."*

This sentiment was echoed in Intibucá:

> *"There are organised groups in the aldeas and they haven't been able to go through with their projects because of the lack of the transfer to the municipalities. So we advocate to [AHMON] so they think about the women in these communities and they see that there is a barrier for the [local] women's organisations."*

2) Was the advocacy work carried out by the project at the local level reflected at the national level? If so, to what extent?

While the women respondents did not specifically mention national-level changes, their response to this question instead stressed the importance of municipal policy.

In particular, when asked about the local project's influence at the national level, women from all three departments specifically responded with the example of influencing the municipal budget. Furthermore, women involved in the project are able

to encourage *other* women to speak out and help through various legal or political processes, such as reporting violence (Lempira), auditing water committees (La Paz), or encouraging young women to participate in career training (Intibuca).

Also of note, women spoke of local-level project activities as emboldening women's government influence. For example, respondents from La Paz mentioned that marches and protests have changed the way government is following up on complaints about violence against women.

> "*Here we have done everything including shouting. We have marched, we have shouted at the public establishments so that [acts of violence] do not go unpunished, and we have achieved it. I think that now, when they get a report [of violence], they follow up on it."*

Finally, the women stated that the linkage between local networks and national networks allows information to flow upward. For example, women from Intibuca mentioned that the use of funds at the local level is recorded and sent to the national organisation. La Paz respondents stated that issues experienced at the local level, such as issues regarding health, violence, or education, are also communicated to the national organisation via regular auditing reports compiled by the women's networks. Court reports on the number of accusations of violence that have been filed are included in these audits, so that the national organisation can discuss various solutions. However, it is unclear from the qualitative evidence whether this upward flow of information has had any impact on the actions of the project at the national level.

3) What are the advantages, disadvantages and challenges in developing projects that work simultaneously at the local and national level?

Responses to this question varied widely between focus groups. Table 7.1 provides a summary, and highlights commonalities.

Table 7.1: advantages, disadvantages, and challenges of developing projects that work at a local and national level, by department

Advantages	Disadvantages	Challenges
INTIBUCA		
More strength, unity, and support	Groups are not consolidated	Need for compromise
Ability to establish more partnerships	Lack of loyalty	***Transparency with funding***
Maintaining the organisation	Inadequate monitoring and follow-up	
Improves communication at all levels	***Poor management of funds***	
More opportunities for assistance		
LA PAZ		
More security, strength, certainty	***Insufficient funding***	
Facilitates management at local level		
Ability to participate at many levels: national, regional, local		
Empowerment through learning		
LEMPIRA		
More strength in unity	Technical personnel are not nearby	
Facilitates use of technology		
Budget and logistic support		
Ability to expand		

Focus group participants in all three departments claimed that simultaneous work at the local and national level allowed for women's networks to feel unified and strengthened. It also allowed for freedom of participation, by involving other networks or expanding. However, concerns about funding were raised in both the Intibuca and La Paz focus groups, with women stating either that funds were being mismanaged or if they were insufficient. Although women from La Paz and Lempira appreciated the management support received from upper-level organisations, one challenge of the multi-level project is ensuring that this transparency works in the opposite direction as well.

8 CONCLUSIONS

8.1 CONCLUSIONS

This evaluation has found evidence that, six months after its conclusion, the project 'Promoting women's engagement in governance in Honduras' had a positive impact on overall women's empowerment.

In order to measure women's empowerment, the evaluation identified and employed 14 indicators describing an empowered woman in the three departments of Honduras, and assessed how women involved in the project compared relative to these indicators with similar women not involved in the project. The project participants on average score positively in 55 per cent of the indicators, compared with 45 per cent for women not involved in the project. This difference is statistically significant, suggesting that the project had a positive impact on overall women's empowerment.

The project appears to have had a positive and significant impact on indicators referring to knowledge of individual rights, willingness to defend rights, and willingness to report violence. There is also evidence to suggest that the project increased women's likelihood of deeming violence against her to be unacceptable, though the results in both intervention and comparison groups are very positive in this respect. The indicator measuring personal autonomy has high values both intervention and comparison groups, and the evaluation finds no significant differences between the two. Finally, there is no evidence of significant impact in terms of other indicators measuring personal empowerment (self-confidence and opinion on women's economic role), and the low values for each suggest room for improvement.

The evaluation also finds positive and significant results of the project on some key indicators of empowerment at the relational level. Project participants were significantly more likely than comparison respondents to have reported influence in group decision making, higher shares of household income, and ability to influence local policies. Indicators measuring decision-making influence in the household and political sphere, however, provided no evidence of impact. Additionally, while women from project communities and comparison communities reported no difference in the attitudes of community members regarding social norms, project participants were more likely to be aware of municipal resources for women, suggesting that the project had some limited positive impact at the environmental level of empowerment.

Finally, the evaluation also investigated additional indicators outside of those used for the Women's Empowerment Index. Evidence suggests that project participants were more willing to report violence to the authorities. It is important to be aware that women in the project were more likely to report having known of violence committed against women close to them, and more likely to have heard of cases of violence being reported to the authorities. However, it is difficult to determine if this is due to an increase in incidence of violence or an increase in awareness of it. Project teams should not ignore the possibility that the project resulted in increased acts of violence

against participants and women close to them. In order to identify any unintended consequences, future projects should include careful monitoring of cases of violence. Regarding political indicators, there is no evidence that political group participation differs between women in project and comparison groups, but project participants were more likely to have attended political events such as marches and protests than women from the comparison group.

Finally, the qualitative component explored the interconnections between national-level influencing and community-level activities. It found some evidence that the national-level activities were instrumental in increasing knowledge of their rights and supported to vocalize their needs and advocate for their rights at local level. Focus group discussions also revealed some up-ward stream of information from local to national, specifically on health, violence, and education. The focus groups also highlighted a need for the national-level activities to be more transparent in their allocation of funds, in order to foster trust and understanding of the use of network resources.

8.2 PROGRAMME LEARNING CONSIDERATIONS

Consider expanding activities promoting women's political influence.

This evaluation finds positive and significant impact on women's ability to influence change, as demonstrated by a range of indicators (for example, knowledge of rights, influence in local policies, and participation in political events) as well as qualitative results. Given the impact of these community-level activities, future projects are encouraged to learn from these successes, so as to encourage women's advocacy in additional community-level policies. These activities could also be broadened in order to achieve policy objectives on a broader geographical level.

Consider revising programme logic to make causal chain for certain goals more explicit.

The evaluation did not provide evidence of positive impact for a few outcomes explicitly stated in the logic model, specifically *self-confidence* and *household decision-making*. However, for contribution to household income – an outcome not as well defined by the theory of change – the project proved very effective. Consider revisiting the logic model, reconsidering assumptions and identifying gaps in the causal chain that may have impeded impact or created unexpected positive change. For example, what additional steps or activities might be necessary to improve self-confidence? What aspects of the project led to an increase in women's contributions to their household income?

Consider altering the targeting strategy.

From the evaluation it emerged that project participants were more likely than the comparison respondents to be already participating in groups and political activities. Additionally, project participants were more highly educated and from wealthier households, on average. The project team is encouraged to explore alternative targeting strategies in order to examine if project activities have even stronger impacts with women who could be considered more vulnerable.

Promote transparency regarding activities and funds used at the national level.

Responses from focus groups highlighted the positive impact of local women's involvement in activities at the national level. Still, some expressed concerns over the

use of national-level funds. While this does not necessarily signal misuse, it does suggest that focusing on transparency of funds at the national level and clearly communicating how these finances are allocated could serve to further strengthen the trust and collaboration between networks.

Further explore how women allocated their share of the municipal budget, as these may have caused unexpected positive impacts.

Results suggest a positive impact on women's contribution to household income. While it is still unclear exactly how the project contributed to this outcome, it is likely that project women effectively used the 5 per cent of the municipal budget to fund new economic activities. An increased understanding of the specifics of these activities may allow future projects to replicate this positive impact.

Explore activities to empower women at the household-level

Programme women displayed higher levels of empowerment in *group decision-making*, an outcome explicitly linked to their involvement in women's networks. However, another indicator focusing on women's sphere of influence in the household, *opinion on women's economic role,* showed no effects. Future projects can build on the successes of this programme by targeting household-level power relations to further empower women in various areas of their lives.

APPENDIX 1: THRESHOLDS FOR CHARACTERISTICS OF WOMEN'S EMPOWERMENT

Level	Characteristic	Threshold: a women scores positively if she...
Personal	Self-confidence	Responds positively to all four of the following questions. That is, she expresses she "completely agrees" with the first two statements and last statement, and "completely disagrees" with the third statement. • I believe I am a person of worth, on an equal plane with others • I believe I have many good qualities or strengths • I do not believe I have much to be proud of • I am equal to my peers (siblings, friends, colleagues, etc.)
	Knowledge of individual rights	• Responds correctly to the question "what percentage of the municipal budget should be invested in actions that benefit women?" The correct answer is 5 per cent.
	Unacceptability of violence	Does *not* consider violence against a woman acceptable in any of the following 7 circumstances: • If she disobeys her husband or other family members. • If he suspects that she has been unfaithful. • If she neglects the children. • If she spends money without permission. • If she is not supporting her husband in livestock and agricultural activities. • If she goes to see her family without the permission of her husband. • Any other case not mentioned above
	Personal autonomy	Confirms she is a decision-maker in both of the following scenarios: • If she wants to travel to visit family that live outside the community • If she participates in community meetings or group activities
	Willingness to defend own rights	Responds positively to the following three statements. That is, she expresses agreement with the first two statements, and disagreement with the third. • Public meetings held in your village are quite intimidating - it is difficult for a woman like you to stand up and voice any concern • If a public decision was made which might negatively affect your life and those of your children, you would not hesitate to stand up and protest despite the possible negative consequences • I do not feel intimidated to go and talk to the major if some of my rights are not respected
	Willingness to report cases of violence	• Responds that "yes," she would be willing to report a man if he beat her (and scores negatively if she answers "no" *or* declines to answer)

Level	Characteristic	Threshold: a women scores positively if she...
	Opinion on women's economic role	Responds positively to both of the following statements. That is, they strongly agree with the first statement, and strongly disagree with the second. • Women are just capable as men of contributing to household income • A man's job is to earn money; a women's job is to look after home and family
Relational	Group decision making	• She expresses she is moderately or very able to influence decisions in at least one community group in which she participates.
	Household decision making	She confirms she is involved in making decisions for all of the following areas relevant to her household. • How cash earned from agricultural production is spent. • Purchase and sale of large animals. • Purchase and sale of land. • Purchase of furniture and electronic devices. • Whether the household should borrow a small loan, from what source, how much to borrow, and how to spend the money. • Whether to migrate, to what place, and when. • What other non-agricultural business activities your household will engage in, and how to manage these activities. • How cash earned from these non-agricultural activities is spent. • What to give relatives when they marry or have a celebration. • The education of your children. • Transfer of property to a relative or any other person.
	Share of household income	• Reports contributing over 50 per cent of the household income, or if her contribution has increased since 2008.
	Influencing local policies	• Responds that "yes," she has tried to influence the way the municipal budget is invested in her community.
	Political decision making	• She expresses she is moderately or very able to influence decisions in at least one political group in which she participates.
Environmental	Social norms	She believes that women from her community would respond positively to the following statements, and/or men from her community would respond positively to all following statements: • Women are as capable as men at contributing to household income • Women can be as good leaders as much as men can • A man's work is to provide income, and a woman's work is to take care of the family and the house
	Awareness of resources for women	• Responds that "yes," she has heard about municipal resources that serve to benefit women.

APPENDIX 2: METHODOLOGY USED FOR PROPENSITY-SCORE MATCHING

The analysis of outcome variables presented in Section 5 of this report involved group mean comparisons using propensity-score matching (PSM). The basic principle of PSM is to match each participant with a non-participant that was observationally similar at baseline and to obtain the treatment effect by averaging the differences in outcomes across the two groups after project completion. Unsurprisingly, there are different approaches to matching, i.e. to determining whether or not a household is observationally 'similar' to another household. For an overview, we refer to Caliendo and Kopeinig (2008).[3] This appendix describes and tests the specific matching procedure followed in this Effectiveness Review.

Estimating propensity scores

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

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

Table A2.1: Estimating the propensity score

	Marginal effect	Standard error	p-value
Household surveyed:			
1[Lempira department]	0.07	0.05	0.23
1[Respondent's ethnic group is Lenca]	-0.45***	0.11	0.00
1[Respondent's religion is Catholic]	0.23***	0.05	0.00
1[Respondent is married]	0.06	0.06	0.30
Respondent's age (years)	-0.00	0.00	0.59
1[Respondent has primary education at least]	0.15	0.08	0.07
1[Respondent is able to work]			
Number of members of the household	0.04**	0.01	0.00
1[Respondent is household head]	0.16*	0.07	0.01
1[Respondent's contribution to total household income >= 50% in 2008]	0.12*	0.06	0.03
1[Household was in the lowest quintile of the wealth index in 2008]	-0.22**	0.07	0.00
1[Household was in the second quintile of the wealth index in 2008]	-0.15*	0.07	0.04
1[Household was in the third quintile of the wealth index in 2008]	-0.05	0.07	0.47
1[Household was in the fourth quintile of the wealth index in 2008]	0.08	0.08	0.32
1[Respondent worked land or raised animals in 2008]	-0.01	0.06	0.93
1[Respondent had formal or temporal employment in 2008]	-0.05	0.08	0.57
1[Respondent participated in a community group in 2008]	-0.03	0.05	0.54
1[Respondent participated in a political group in 2008]	0.15**	0.06	0.01
Observations	489		

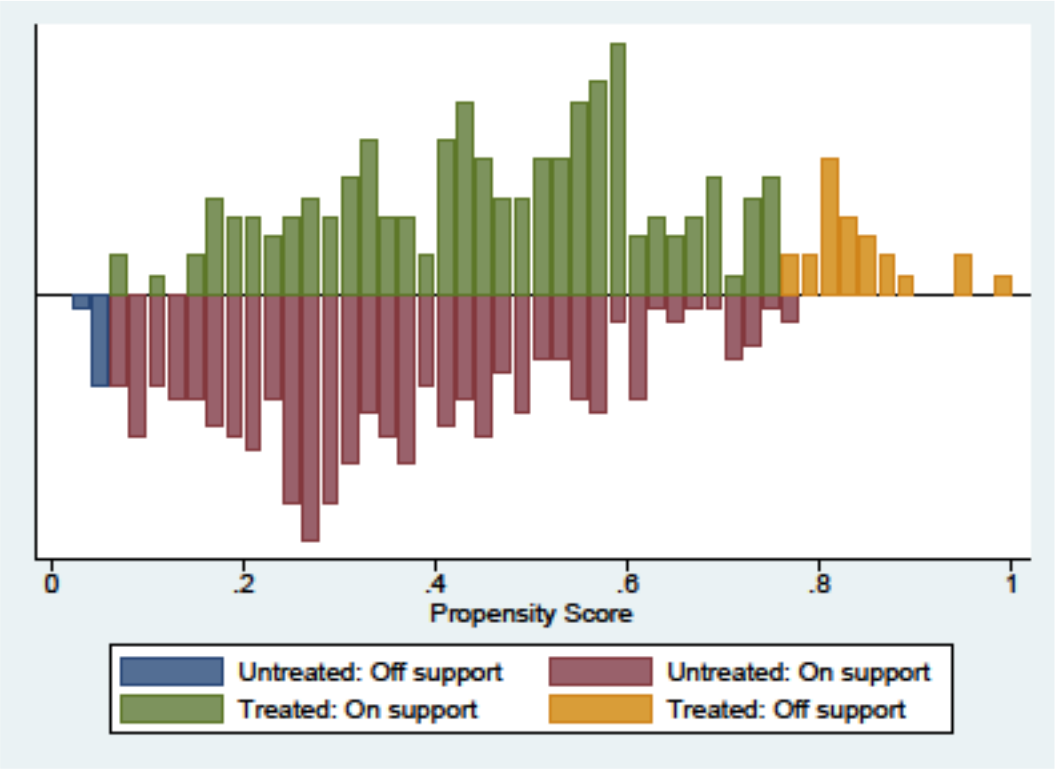
Marginal effects. The construction of the wealth index is described in Section 4.2. Variables dated 2008 are estimates, based on recall data. Dependent variable is binary, taking 1 for project participant households, and 0 otherwise. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Defining the region of common support

After estimating the propensity scores, the presence of a good *common support area* needs to be checked. The area of common support is the region where the propensity score distributions of the treatment and comparison groups overlap. The common support assumption ensures that 'treatment observation have a comparison observation "nearby" in the propensity score distribution' (Heckman, LaLonde and Smith, 1999). Since some significant differences were found between the intervention and comparison groups in terms of their baseline characteristics (as detailed in Section 4.2), some of the women in the intervention group were too different from the comparison group to allow for meaningful comparison. We used a minima and maxima comparison, deleting all observations whose propensity score is smaller than the minimum and larger than the maximum in the opposite group (Caliendo and Kopeinig, 2008). Of the 200 women interviewed in the project communities, 27 were dropped because they lay outside the area of common support. Therefore, the results of the evaluation can only be interpreted with reference to the remaining 173 women, and is not necessarily representative of the impact on all project participants.

Figure A2.1 illustrates the propensity scores and shows the proportion of women lying on and off the areas of common support, by treatment group.

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



Matching intervention and comparison households

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

The *psmatch2* module in Stata[4] was used with a bandwidth of 0.06 and with the analysis restricted to the area of common support.

When using PSM, standard errors of the estimates were bootstrapped using 1,000 repetitions, to account for the additional variation caused by the estimation of the propensity scores and the determination of the common support.[5]

Check balancing

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

Table A2.2: Balancing test

Variable	Unmatched Matched	Treated Mean	Comparison mean	% bias	% reduction bias	t	p>|t|	V_e(T) / V_e(C)
1[Lempira department]	U	.69364	.69014	0.8		0.08	0.938	1.00
	M	.69364	.68674	1.5	-97.1	0.14	0.890	1.00
1[Respondent's ethnic group is Lenca]	U	.98266	.98944	-5.8		-	0.538	1.65*
	M	.98266	.97688	4.9	14.7	0.38	0.703	0.95
1[Respondent's religion is Catholic]	U	.76879	.60211	36.4		3.71	0.000	0.79*
	M	.76879	.75548	2.9	92.0	0.29	0.772	1.03
1[Respondent is married]	U	.45087	.40141	10.0		1.04	0.300	1.03
	M	.45087	.44765	0.7	93.5	0.06	0.952	1.00
Respondent's age (years)	U	40.017	38.176	14.9		1.54	0.124	0.97
	M	40.017	40.189	-1.4	90.7	-	0.897	0.89
1[Respondent has primary education at least]	U	.12717	.09859	9.0		0.95	0.344	1.26*
	M	.12717	.12758	-0.1	98.5	-	0.991	1.00
1[Respondent is able to work]	U	1	1	.		.	.	.
	M	1	1	.	.	.	.	.
Number of members of the household	U	5.3931	5.1796	10.0		1.04	0.297	1.15
	M	5.3931	5.3566	1.7	82.9	0.16	0.876	1.10
1[Respondent is household head]	U	.30636	.25704	11.0		1.14	0.253	1.17
	M	.30636	.31501	-1.9	82.5	-	0.862	0.96
1[Respondent's contribution to total household income >= 50%	U	.31214	.22535	19.6		2.06	0.040	1.24
	M	.31214	.31821	-1.4	93.0	-	0.904	1.01
1[Household was in the lowest quintile of the wealth index in	U	.13873	.24648	-		-	0.006	0.63*
	M	.13873	.14462	-1.5	94.5	-	0.876	0.95
	U	.17341	.21479	-		-	0.284	0.84

1[Household was in the second	M	.17341	.17735	-1.0	90.5	-	0.924	0.97
1[Household was in the third quintile of the wealth index in	U	.23121	.19366	9.2		0.96	0.338	1.17
	M	.23121	.22486	1.6	83.1	0.14	0.888	1.04
1[Household was in the fourth quintile of the wealth index in	U	.21965	.16549	13.7		1.44	0.150	1.22
	M	.21965	.23143	-3.0	78.3	-	0.794	1.01
1[Respondent worked land or raised animals in 2008]	U	.72254	.72535	-0.6		-	0.948	1.01
	M	.72254	.74231	-4.4	-603.7	-	0.679	0.87
1[Respondent had formal or temporal employment in 2008]	U	.10405	.09155	4.2		0.44	0.661	1.14
	M	.10405	.11127	-2.4	42.2	-	0.829	1.02
1[Respondent participated in a community group in 2008]	U	.56647	.50352	12.6		1.31	0.192	0.95
	M	.56647	.56083	1.1	91.0	0.11	0.916	1.01
1[Respondent participated in a political group in 2008]	U	.34682	.20423	32.3		3.41	0.001	1.37*
	M	.34682	.33902	1.8	94.5	0.15	0.879	1.00

* if 'of concern', i.e. variance ratio in [0.5, 0.8) or (1.25, 2]
** if 'bad', i.e. variance ratio <0.5 or >2

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	% concern	% bad
Unmatched	0.080	48.71	0.000	13.4	10.5	69.8*	0.85	28	0
Matched	0.002	0.75	1.000	2.0	1.6	9.3	0.84	0	0

* if B>25%, R outside [0.5; 2]

APPENDIX 3: ROBUSTNESS CHECKS

In order to address the validity of the results presented in Section 5, a series of robustness checks were carried out to check if the preferred matching algorithm is the one that best performs the matching between intervention and comparison groups. This section presents a number of alternative matching algorithms used to test the robustness of the estimates presented in Section 5.

1. Multivariate regression

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

$$Y_i = \alpha + \beta_1 Project\ participation_i + \delta^{'} X_i + \varepsilon_i$$

Where Y_i is the dependent variable; X_i is a vector of household covariates used in the model in table A2.1; finally the variable of interest is the dummy variable *Project Participation* that assumes value equal to one when the household is enrolled in the project, zero otherwise. When the dependent variable Y_i is binary variable, a probit model replaces the OLS specification. It is important to note that in the absence of randomized allocation of the project among the population in our sample, OLS and probit models fail to identify the causal effect of the programme, and can only account for observable differences between intervention and comparison groups. Only the estimate of β_1 will be reported.

2. Propensity Score Weighting

Following the example of Hirano and Imbens (2001)[6] we implemented a regression adjustment with weights based on the propensity score. The average treatment effect can be estimated in a parametric framework as follows:

$$Y_i = \alpha + \beta_1 Project\ participation_i + \delta_2' Z_i + \delta_1' X_i + \varepsilon_i$$

Where Y_i represents the outcome of interest; $\mathrm{Project\ participation_i}$ is a dummy binary variable equal to one if an individual/household is enrolled into the programme and zero otherwise; X_i is a vector of matching covariates used to estimate the propensity score match; and Z_i is a vector of control variables which cannot be used for the matching as they are not supposed to influence project participation. The regression is estimated with weights equal to one for the treated units and $\hat{e}(x)/(1-\hat{e}(x))$ for control units. This parametric regression analysis framework has the advantage of exploring heterogeneity in the treatment effect. Moreover it allows controlling for variables that cannot be included in the propensity score equation. The robustness check tables will only report β_1.

3. Propensity Score Matching – Nearest Neighbour

The Nearest Neighbour (NN) matching algorithm finds an observation from the comparison group to be matched with an observation from a treated individual that is closest in terms of their propensity score.[7] In this robustness check we apply the NN method estimating the average treatment effect on the treated, with one match per observation, using Abaide-Imbens standard deviations and exact match on the partner organisation.

Table A3.1: Overall women's empowerment index

	1	2	3

	Women's Empowerment Index (Personal)	Women's Empowerment Index (Relational)	Women's Empowerment Index (Enviromental)
OLS	0.05***	0.04**	0.20***
	(0.01)	(0.01)	(0.01)
N	799	799	799
Propensity Score Weighting	0.05***	0.04***	0.20***
	(0.01)	(0.01)	(0.01)
N	796	796	796
Nearest neighbour	0.08***	0.03*	0.21***
	(0.02)	(0.02)	(0.02)
N	796	796	796

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

	1 Women's Empowerment Index
OLS	0.11***
	(0.02)
N	492
Propensity Score Weighting	0.10***
	(0.02)
N	457
Nearest neighbor	0.09***
	(0.03)
N	457

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.2: Women's empowerment – power dimensions

	1 Empowerment base index (personal)	2 Empowerment base index (relational)	3 Empowerment base index (environmental)
OLS	0.11***	0.10***	0.12***
	(0.02)	(0.02)	(0.03)
N	492	492	492
Propensity Score Weighting	0.10***	0.10***	0.12***
	(0.02)	(0.02)	(0.03)
N	457	457	457
Nearest neighbor	0.11***	0.10***	0.12***
	(0.02)	(0.02)	(0.03)
N	492	492	492

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.3: Personal-level empowerment

	1 Self-confidence	2 Personal autonomy	3 Knowledge of women's rights	4 Willingness to influence and defend own rights	Opinions on women's economic role
OLS	0.00	0.02	0.34***	0.15***	0.00
	(0.05)	(0.02)	(0.04)	(0.05)	(0.05)

N	492	492	492	492	492
Propensity Score Weighting	-0.02	0.01	0.35***	0.13***	-0.02
	(0.05)	(0.03)	(0.05)	(0.05)	(0.05)
N	457	457	457	457	457
Nearest neighbor	0.00	0.02	0.34***	0.15***	0.00
	(0.05)	(0.02)	(0.04)	(0.05)	(0.05)
N	492	492	492	492	492

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.4: Personal-level empowerment (cont)

	1 Unacceptability of gender based violence	2 Willingness to report violence
OLS	0.10***	0.10**
	(0.03)	(0.04)
N	492	492
Propensity Score Weighting	0.10***	0.12***
	(0.04)	(0.04)
N	457	457
Nearest neighbor	0.10***	0.10**
	(0.03)	(0.04)
N	492	492

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.5: Relational-level empowerment (power over)

	1 Household decision making [proportion of activities]	2 Share of household income [proportion of resources]
OLS	0.04	0.15***
	(0.05)	(0.04)
N	492	492
Propensity Score Weighting	0.03	0.15***
	(0.05)	(0.04)
N	457	457
Nearest neighbor	0.04	0.15***
	(0.05)	(0.04)
N	492	492

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.6: Relational-level empowerment (power with)

	1 Group decision making	2 Influence in local policies	3 Political decision making
OLS	0.19***	0.13***	-0.01
	(0.05)	(0.04)	(0.04)
N	492	492	492
Propensity Score Weighting	0.19***	0.14***	-0.02
	(0.05)	(0.04)	(0.04)
N	457	457	457
Nearest neighbor	0.18**	0.12*	0.01
	(0.08)	(0.06)	(0.07)
N	457	457	457

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.7: Environmental-level empowerment

	1 Social Norms	2 Awareness municipal resources for women
OLS	0.05	0.18***
	(0.04)	(0.05)
N	492	492
Propensity Score Weighting	0.06	0.17***
	(0.04)	(0.05)
N	457	457
Nearest neighbor	-0.01	0.14**
	(0.07)	(0.07)
N	457	457

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.8: Exposure to violence in the past 12 months

	1 Exposure to violence (respondent)	2 Exposure to violence (women close to respondent)
OLS	0.04	0.16***
	(0.05)	(0.05)
N	415	415
Propensity Score Weighting	0.03	0.13**
	(0.05)	(0.06)
N	383	383
Nearest neighbor	0.12	0.15*
	(0.08)	(0.08)
N	383	383

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.9: Types of violence experienced in the past 12 months

	1 Type of violence: psychological	2 Type of violence: physical	3 Type of violence: sexual
OLS	0.15***	0.04	0.04
	(0.05)	(0.05)	(0.03)
N	492	414	492
Propensity Score Weighting	0.14***	0.04	0.04
	(0.05)	(0.05)	(0.03)
N	457	382	457
Nearest neighbor	0.13*	0.00	0.04
	(0.07)	(0.07)	(0.05)
N	457	382	457

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.10: Reported violence to an authority in the last 3 years

	1 Reported violence

OLS	0.10**
	(0.04)
N	491
Propensity Score Weighting	0.11**
	(0.05)
N	456
Nearest neighbor	0.10
	(0.06)
N	456

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3.11: Political influence and participation

	1 Started political participation	1 Event participation	1 Number of events participated in
OLS	0.07*	0.15***	3.64***
	(0.04)	(0.04)	(0.77)
N	492	492	492
Propensity Score Weighting	0.06	0.17***	3.71***
	(0.04)	(0.04)	(1.05)
N	457	457	457
Nearest neighbor	0.06	0.20***	4.18***
	(0.06)	(0.06)	(0.92)
N	457	457	457

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.

NOTES

[1] Lombardini, Simone, Kimerbly Bowman, and Rosa Garwood. 2017. "A 'How To' Guide to Measuring Women's Empowerment: Sharing experience from Oxfam's impact evaluations" Oxford: Oxfam GB.

[2] Estimates from the robustness checks, reported in Appendix 3, do not provide strong evidence in either direction. With the OLS model and Propensity Score Weighting providing similar effect in magnitudes and statistically significant at 1% level, and PSM with nearest neighbour providing smaller estimates not statistically significant.

3 Caliendo, M. and Kopeinig, S. 2008. Some Practical Guidance for the Implementation of Propensity Score Matching, *Journal of Economic Surveys*, Wiley Blackwell, vol. 22(1), pages 31–72.

[4] E. Leuven and B. Sianesi. (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html.

5 Bootstrapping is a statistical procedure where repeated samples are drawn from the original sample with replacement. This results in a statistical distribution of parameter estimates (the sampling distribution). The bootstrapped standard error is the standard deviation of this sampling distribution and it can be shown that as the number of repeated samples becomes large, provided certain technical conditions are met, this is a good estimate for the standard error of the estimate.

6 Hirano, K. and Imbens G.W. (2001), Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services & Outcomes Research Methodology*, vol. 2, pp. 259–278.

[7] Several variants of NN matching are possible, e.g. NN matching 'with replacement' and 'without replacement'. In the former case, an untreated individual can be used more than once as a match, whereas in the latter case it is considered only once. Matching with replacement involves a trade-off between bias and variance. If we allow replacement, the average quality of matching will increase and the bias will decrease. This is of particular interest with data where the propensity score distribution is very different in the treatment group and the control group (Caliendo and Kopeinig, 2008).

Oxfam Effectiveness Reviews

For more information, or to comment on this report, email ppat@oxfam.org.uk

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

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

OXFAM

Oxfam is an international confederation of 17 organisations networked together in more than 90 countries, as part of a global movement for change, to build a future free from the injustice of poverty:

Oxfam America (www.oxfamamerica.org)
Oxfam Australia (www.oxfam.org.au)
Oxfam-in-Belgium (www.oxfamsol.be)
Oxfam Canada (www.oxfam.ca)
Oxfam France (www.oxfamfrance.org)
Oxfam Germany (www.oxfam.de)
Oxfam GB (www.oxfam.org.uk)
Oxfam Hong Kong (www.oxfam.org.hk)
Oxfam India (www.oxfamindia.org)
Intermón Oxfam (Spain) (www.intermonoxfam.org)
Oxfam Ireland (www.oxfamireland.org)
Oxfam Italy (www.oxfamitalia.org)
Oxfam Japan (www.oxfam.jp)
Oxfam Mexico (www.oxfammexico.org)
Oxfam New Zealand (www.oxfam.org.nz)
Oxfam Novib (www.oxfamnovib.nl)
Oxfam Québec (www.oxfam.qc.ca)

Please write to any of the agencies for further information, or visit www.oxfam.org.