EVALUATION: JANUARY 2016 PUBLICATION: NOVEMBER 2017

RESILIENCE IN NEPAL

Impact evaluation of the Joint Programme on Disaster Risk Management and Humanitarian Preparedness

Effectiveness Review Series 2015/16

Photo: Women's empowerment groups are adopting a system of rice intensification method in Sarlahi District. Credit: Bagmati Welfare Society Nepal (BWSN)

JONATHAN LAIN

OXFAM GB

www.oxfam.org.uk/effectiveness

OXFAM

ACKNOWLEDGEMENTS

We would like to thank the staff of the partner organisations and of Oxfam in Nepal for their support in carrying out this Effectiveness Review. Particular thanks are due to Basu Neure, Sumeru Tripathee and Nidhi Shrivastav. Thanks are also due to Kiran Chalise and Madhu Sudan Gautam for their excellent leadership of the survey process. Finally, we thank Rob Fuller for his vital comments on earlier drafts of this report, and Karuna Amatya and Gaurab Sagar Dawadi for their additional feedback.

CONTENTS

Contents 3

Executive Summary 4

1 Introduction 10

2 Project description 12

3 Evaluation design 16

4 Data 19

5 Measuring Resilience in Nepal 25

6 Results 31

7 Conclusions 56

EXECUTIVE SUMMARY

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

During the 2015/16 financial year, one of the projects that was randomly selected for an Effectiveness Review was the 'Joint Programme on Disaster Risk Management and Humanitarian Preparedness'. This project was carried out in Nepal by Oxfam in partnership with several organisations, including the Koshi Victims Society (KVS), the Social Development Research Centre (SDRC), Bagmati Welfare Society Nepal (BWSN), Nepal Red Cross Society (NRCS), and Rural Development Centre (RDC). The project activities, which began in April 2011 and finished in March 2016, were implemented in four districts in the Terai region of southern Nepal – Dhanusha, Rautahat, Salarhi, and Saptari.

The project was designed to build the resilience of project participants to a number of different shocks and stresses. Households' livelihoods in the region are typically based around agricultural activities, which are threatened by floods, droughts, outbreaks of human and animal diseases, and spells of especially cold weather. The project worked at a number of different levels to reduce households' vulnerability to these risks. Within communities, a programme of training and support was provided to enhance the activities of local disaster management committees and women's empowerment groups (WEGs). Improved Water, Sanitation and Hygiene (WASH) facilities were also constructed, as well as small-scale flood mitigation structures, such as embankments, culverts, and dams. The project also supported district- and national-level institutions to help create disaster management plans, which should have a wider impact on resilience. Furthermore, the project undertook some advocacy activities to try and change national policies related to disaster management.

Evaluation approach

This Effectiveness Review used a quasi-experimental evaluation design to assess the impact of the project activities among the households whose members directly participated in the women's empowerment groups that were formed by the project, and through which many of the project activities were channelled. This involved comparing those households that participated in the project to a group of comparison households, which were similar to the project participants. This means that this Effectiveness Review can only identify the household- and community-level effects of the project. Activities operating at the district or national level, including the project's advocacy work, are not included in this evaluation.

In the project communities, 280 project households from eight Village Development Committees (VDCs) – which are similar to municipalities – were randomly sampled from household lists that were created using lists of WEG members. For the comparison, 12 non-project VDCs were identified, which were similar to the project communities in terms of a number of key characteristics, including the dominant livelihood strategies employed by community members, the distance of the community from large rivers, and the ethnic composition of the community. Within comparison VDCs, a list of households was created using voter lists obtained from the Electoral Commission, and 529 respondents were randomly sampled. Households that were too wealthy or too highly educated to be eligible for participation in the project activities

were systematically excluded from the sample using screening questions asked at the start of the questionnaire.

At the analysis stage, the statistical tools of propensity-score matching and multivariate regression were used to control for apparent baseline differences between the households in the project and comparison communities, to increase confidence when making estimates of the project's impact.

The primary aim of the Effectiveness Review was to investigate the project's impact on building resilience to shocks and stresses. This was assessed by identifying 33 characteristics that are thought to be associated with resilience, for which data could be collected in the household survey. These characteristics were formulated under Oxfam GB's multidimensional framework for measuring resilience, and developed through discussions with project staff and focus groups conducted in local communities. In general, these indicators were chosen to focus on the intermediate steps between project activities and final well-being outcomes. A full list of indicators and a summary of the results for each is shown in Table 1.

Table 1: Characteristics of resilience examined in this Effectiveness Review

Dimension	Characteristic	Connected to project logic?	Evidence of positive impact?
Livelihood viability	Ownership of productive assets	Yes	No
	Ownership of land	No	Yes
	Ownership of livestock	Yes	Yes
	Livelihood diversification	Yes	No
	Crop diversification	Yes	Yes
	Cultivation of drought-resistant crops	Yes	Yes
	Dietary diversity.	No	Yes
	WASH attitudes	Yes	No
	Livestock hygiene practices	No	Yes
Innovation potential	Attitude to change	No	No
	Access to credit	Yes	Yes
	Awareness of climate change	No	No
	Adoption of innovative practices	No	No
	Access to markets (in monsoon season)	Yes	No
Access to contingency resources and support	Participation in community groups	Yes	Yes
	Social connectivity	Yes	Yes
	Savings	Yes	Not clear
	Formal earnings	No	No
	Awareness of community disaster management plan	Yes	Yes
	Awareness of VDC disaster management plan	Yes	Yes
	Awareness of district management plan	Yes	Yes
	Direct access to emergency items	Yes	Yes
Integrity of natural and built environment	Availability of drinking water	Yes	No
	Irrigation for agriculture	Yes	Yes
	Protection of house from flood	No	Yes
	Protection of house from fire	No	No
	Use of organic fertiliser	Yes	No
	Caring for forest/greenery	No	Yes
	Participation in disaster planning process	Yes	Yes

Social and institutional capability	Early-warning system	Yes	Yes
	Awareness of local leaders' planning activities	Yes	Yes
	Women's representation in the community	Yes	Yes
	Registration of community groups	Yes	Yes

Results

Our data suggest that the project improved the resilience of project households substantially. Project households scored positively on average in 54 percent of the resilience indicators identified for this Effectiveness Review, compared with just 36 percent for the comparison group. It is, however, important to note that many of the resilience indicators for which project households outperformed comparison households were 'output-related', as they were directly linked to the project activities. For example, some of the largest impacts of the project were seen in terms of indicators related to the disaster-planning process.

There were, however, some higher-level indicators of resilience for which our data suggest the project had a positive impact. Project households had better access to credit, were taking more action to protect local natural resources (such as forests), and had improved dietary diversity compared with their non-project comparators. This highlights the importance of considering each resilience indicator separately, as well as the aggregated indices of resilience.

There were a number of other aspects of the project logic at earlier stages of the causal chain, which it was possible to investigate in this Effectiveness Review. Considering agricultural activities first, the project appeared to have positive effects on the rate of livestock ownership and the number of types of animals that households owned. These effects on the portfolio of livestock owned appear to have been concentrated among large animals, such as cows, buffalos, and oxen. However, there was not strong evidence that women's responsibility for households' animals was affected by the project.

While the project did not affect households' decisions over whether or not to engage in farming, there were positive and significant effects on the amount of land that households farmed and the number of crops in their portfolio. The project had especially large positive impacts on the cultivation of drought-resistant varieties of rice and wheat and farming crops in the off-season, and project households were also more likely to sell their agricultural produce. Vitally, women's responsibility for growing crops was found to be moderately higher among project households than those in the comparison group. Nevertheless, there were no clear effects on the use of organic fertilisers, despite the project having provided training and inputs for the use of vermi-compost.

In contrast to the results for agricultural activities, the project did not have clear positive effects on households' engagement in non-farm income-generating activities. If anything, project households were less likely to engage in non-farm livelihood strategies than comparison households, although these results were largely driven by a reduction in the proportion of households engaging in casual labour, such as construction or working short-term in a factory. The main puzzle emanating from these results is that the project did not increase the proportion of households that were operating their own businesses, especially as credit and savings constraints were, in fact, relaxed among the project households. This may be because direct support for new types of livelihood activities was not built into the project until its later stages. However, this may also reflect the fact that the project was successful in supporting agricultural productivity, so non-farm income-generating activities were relatively less lucrative.

Although this Effectiveness Review was focused on resilience, we included some indicators to evaluate the impact on WASH behaviours, given their importance for the project logic. However, the project's apparent effects on WASH were mixed. Although project households had better livestock storage practices, they relied more than the comparison group in our sample on risky water sources – such as ponds, streams, and rivers – for drinking. However, it is important to note that these patterns in drinking water sources were present before the start of the project, so it is very unlikely that the negative differences between the project and comparison households resulted from project activities. Rather, the project was unable to offset the pre-existing differences between project and non-project VDCs.

The main results of this Effectiveness Review are summarised in Table 2. Given the number of indicators that were directly connected to project outputs included in the overall measures of resilience (and its dimensions) it is recommended to look carefully at each constituent indicator when interpreting the results.

Table 2: Key results of this Effectiveness Review

Outcome area		Connected to project logic?	Evidence of positive impact?	Comments
Resilience	Livelihood viability	See Table 1	Yes	The main evidence of impact was seen in ownership of livestock, crop diversification, and dietary diversity. Diversification into non-farm activities was not affected by the project.
	Innovation potential		No	Project households had better access to credit, but their access to markets was more restricted than the comparison households. There was little evidence of any impact on attitudes to change and innovation.
	Access to contingency resources and support		Yes	The strongest impacts were around awareness of disaster management plans at community, VDC and district levels. Group participation and social connectivity were also improved among project households.
	Integrity of natural and built environment		Yes	Project households were more likely to be raised up for flood protection, and care better for local forest/greenery. However, availability of drinking water was actually worse for project households, but this was also the case before the project started, according to recalled data. The project was unable to offset these pre-existing differences between intervention and comparison villages.
	Social and institutional capability		Yes	Many indicators were directly related to project outputs, therefore positive impacts on confidence in early-warning systems, and awareness of/participation in disaster planning are not surprising.
Wealth		No	No	No evidence of impact was found for either current wealth or change in wealth levels during the course of the project.
Livestock		Yes	Yes	Improvements in the livestock portfolio were concentrated in large animals, such as cows, buffalos, and oxen. Women's responsibility for livestock care was unchanged.
Crops		Yes	Yes	Project households had more diversified crop portfolios, and women's responsibility for cultivation increased. There were no clear effects on use of vermi-compost.
Non-farm livelihoods		Yes	No	No changes were identified in terms of household business ownership, while non-

			project households engaged more in off-farm casual labour.
WASH	Yes	Mixed	Project households had improved livestock storage practices, but the initial differences between project and comparison households in terms of drinking water availability were not offset by the project.

Programme learning considerations

Take a more holistic approach to evaluation design, including discussions around suitable indicators and the possible establishment of a comparison group, at the start of future projects.

The quasi-experimental methodology deployed in this Effectiveness Review was successful due to the extensive inputs of project and partner staff during the fieldwork. The discussions around suitable comparison VDCs and the selection of good indicators of resilience were particularly rigorous and inclusive. In part, this was because the set-up for the Effectiveness Review coincided with a monitoring visit, so it was possible to involve a number of representatives from all of the partner organisations in the process. However, it would have been better to have these discussions before the project activities began. As a minimum, this would have helped create a more comprehensive Monitoring, Evaluation, and Learning plan with a wide range of resilience indicators. Going further, it would also have been possible to implement a more robust evaluation design if the comparison group had been established before the project started and baseline data had been collected in project and non-project VDCs. To support this process in the future, the impact evaluation team should provide tools and resources to help project and programme staff with quasi-experimental evaluation designs, especially on setting up a comparison group. Therefore, this learning point applies not only for programme staff, but also those working on monitoring and evaluation.

Build the support for new livelihood activities into the project at its inception rather than during the later stages of its implementation.

The results suggest that the project had far stronger effects on indicators associated with disaster risk reduction than livelihoods. In particular, there were no clear positive effects on project households' likelihood of engaging in non-farm income-generating activities, and there was only limited evidence that project households had a greater propensity to cultivate crops with new technologies, such as vermi-compost. One possible explanation for this is that many of the activities designed to support new livelihoods were only implemented in the latter half of the project's lifetime. Integrating these activities into future projects from the inception phase may increase the chances of effecting greater change in households' livelihood strategies.

Focus more directly on promoting project households' access to markets to encourage further innovation and adaptation.

Overall, there was limited evidence that the project generated clear positive effects on households' potential to innovate. This was despite project households reporting that their access to credit had improved. This may be because project households had, if anything, less access to markets where they could sell their produce than the comparison households in our sample. The cooperatives set up by the project could potentially remedy this issue, so they should receive special attention to ensure households can sell their crops or off-farm goods. It may also be possible to link these cooperatives to formal banking institutions, to further relax the credit constraints faced by households and, potentially, make the borrowing required to facilitate new adaptations even less costly.

Consider avenues for scaling-up the project's successes around local disaster risk reduction plans to other communities.

The most sizeable changes that were generated by the project related to households' awareness of and confidence in disaster management plans at community, VDC and district levels. However, in the non-project communities, awareness of the content of these plans – even district-level plans – was strikingly low. Despite not being the focus of this evaluation, these results suggest that the district level-activities of the project, such as supporting the planning activities of district governments, were only affecting project communities and not the non-project communities in the sample. It is important to find ways of building awareness in VDCs across the district as a whole to ensure these types of district- and national-level advocacy activities are successful.

1 INTRODUCTION

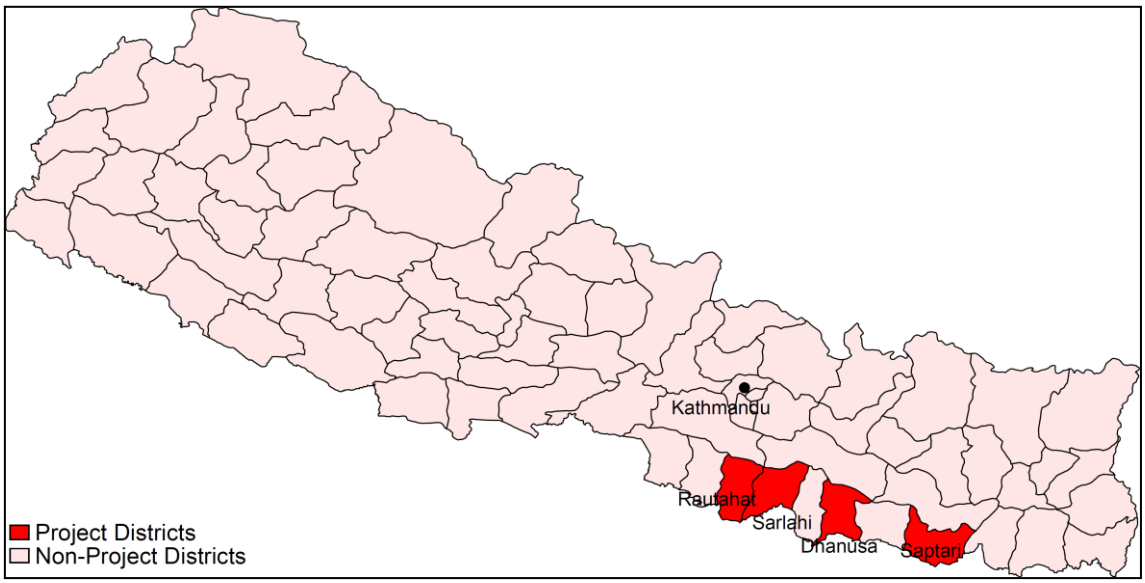
Oxfam GB's Global Performance Framework is part of the organisation's effort to better understand and communicate its effectiveness, as well as enhance learning across the organisation. Under this Framework, a small number of completed or mature projects are selected at random each year for an evaluation 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 Oxfam GB global outcome indicators.

During the 2015/16 financial year, one of the projects that was randomly selected for an Effectiveness Review was the 'Joint Programme on Disaster Risk Management and Humanitarian Preparedness'. This project is being carried out in Nepal by Oxfam in partnership with several organisations, including the Koshi Victims Society (KVS), the Social Development Research Centre (SDRC), Bagmati Welfare Society Nepal (BWSN), Nepal Red Cross Society (NRCS), and Rural Development Centre (RDC). The project activities, which began in April 2011 and finished in March 2016, were implemented in four districts in the Terai region of southern Nepal – Dhanusha, Rautahat, Salarhi, and Saptari (see Figure 1.1).

The project was designed to build the resilience of project participants to a number of different shocks and stresses. Households' livelihoods in the region are typically based around agricultural activities, which are threatened by floods, droughts, outbreaks of human and animal diseases, and spells of especially cold weather. The project worked at a number of different levels to reduce households' vulnerability to these risks. Within communities, a programme of training and support was provided to enhance the activities of local disaster management committees and women's empowerment groups. Improved Water Sanitation and Hygiene (WASH) facilities were also constructed, as well as small-scale flood mitigation structures, such as embankments, culverts, and dams. The project also supported district- and national-level institutions, to help create disaster management plans, which would have a wider impact on resilience. Furthermore, the project undertook some advocacy activities to try and change national policies related to disaster management.

The Effectiveness Review, for which the fieldwork was carried out in January 2016, was aimed at evaluating the success of this project in enabling households to maintain and improve their well-being, in spite of shocks, stresses, and uncertainty. The survey work covered 20 Village Development Committees (VDCs) across all four districts where the project was operating. VDCs correspond to specific geographical areas, similar to municipalities, which are one administrative level below the district.

Figure 1.1 – Project Areas for Nepal Effectiveness Review



Map created by report author (Jonathan Lain), 2016

This report presents the findings of the Effectiveness Review. Section 2 briefly reviews the activities and the logic of the project. Section 3 describes the evaluation design used, and Section 4 describes how this design was implemented. Section 5 presents the approach used to measure resilience. Section 6 shows the results of the data analysis, based on the comparison of outcome measures between project and non-project households. Section 7 concludes with a summary of the findings and some considerations for future learning.

2 PROJECT DESCRIPTION

2.1 PROJECT ACTIVITIES

The project under review worked directly with households to make their livelihoods more resilient to a wide variety of shocks and stresses, while also supporting local-, district-, and national-level institutions to better manage potential risks and disasters.

The project operated mainly in four districts in the Terai region of southern Nepal – Dhanusha, Rautahat, Salarhi, and Saptari. The Terai region comprises a mixture of marshy grasslands, savannahs, and forests, stretching across the flat and relatively low-lying land, south of the outer foothills of the Himalayas. Households in this area typically rely on agricultural activities, based on growing cereals and vegetables that can be consumed by household members or sold to the market. Many households also supplement their agricultural livelihoods by rearing livestock.

These livelihood activities are, however, increasingly under threat from a range of natural and anthropogenic hazards. Most notably, the intensity and timing of seasonal patterns in rainfall can result in flash floods – if the monsoon rains coming in the summer months are too heavy – and droughts – if the monsoon rains come too late. In 2008, rural communities across the entire Terai region were submerged when the embankments of the Koshi River burst, redirecting floodwater across the area of the Koshi alluvial fan. Since 2008, a number of more localised flash floods have affected specific communities within the project area. In 2013, for example, a number of VDCs across Rautahat and Saptari districts were threatened by high water levels in the Bagmati, Lalbakaia and Koshi Rivers. Conversely, during the fieldwork for this review in early 2016, the Terai region was suffering from drought due to the late arrival of the monsoon rains, which may partially be linked to the 2014–16 El Niño.

Communities in the Terai region are also prone to a number of other shocks and stresses, which are in part linked to the sequence of floods and droughts described above. Firstly, the winter months are normally characterised by much cooler weather than the rest of the year, but sometimes the temperature can drop low enough to threaten the health of household members and their animals, during a so-called ‘cold wave’. Households are also exposed to the risk of fires during the dry season, given the proximity of houses to one another and the use of flammable building materials. Finally, flooding may be accompanied by the outbreak of water-borne human and animal diseases.

The 2015 Nepal earthquake also affected some communities in the Terai region, but its impacts were far less severe than areas in and around Kathmandu.

The Joint Programme on Disaster Risk Management and Humanitarian Preparedness (henceforth the ‘Joint Programme’) aimed to build households’ resilience to the shocks and stresses described above, working directly with local communities, as well as district- and national-level authorities. Some project activities had been implemented by Oxfam in collaboration with the local partners as far back as 2008, in direct response to the widespread flooding that affected the Terai region that year, and Oxfam has continued working in the same VDCs until the present day. However, this review focuses solely on the activities of the Joint Programme, which began in April 2011 and finished in March 2016, two months after the fieldwork for this evaluation.

The Joint Programme set three broad objectives during its planning phase:

1. Strengthen and institutionalise Community-Based Disaster Risk Reduction (CBDRR).
2. Enhance the capacity of local institutions to prepare for and respond to humanitarian emergencies.
3. Create an enabling environment for people to demand their 'rights in crisis' by supporting national-level policies, strategies, and guidelines related to disaster management and disaster risk reduction.

Since the project objectives were targeted at community, district and national levels, the project activities were also undertaken at a number of different scales. Firstly, many activities were channelled through groups and civil society organisations within communities, which were either created or supported by the project. The project established over 200 Women's Empowerment Groups (WEGs), which provided training and direct inputs to help members with new flood- and drought-resistant techniques for growing crops, as well as non-farm livelihood strategies. For example, some WEGs were provided with seeds for kitchen gardening, packaging machines for bringing produce to market, and various trainings and exchange visits on the use of organic fertilisers, the management of livestock, and women's leadership. However, it should be noted that these elements of direct support for new or improved livelihood activities only featured in the later stages of the project's implementation. The project also supported the administrative and financial aspects of the WEGs, providing basic stationery and training on record-keeping, as well as mobilising revolving funds to allow WEG members to save and to provide access to credit for new income-generating activities.

The project also channelled a series of activities through pre-existing local groups that worked on managing risks, including Community Disaster Management Committees (CDMCs) and local WASH committees. CDMC members were trained in first aid, search and rescue, and creating local disaster management plans. The project also helped to stockpile resources to enhance early warning systems and improve emergency response, including megaphones, shovels, and buckets. In addition, WASH trainings were organised in project communities to promote good hygiene practices, especially during floods. Vitally, the project supported local groups, such as WEGs, CDMCs and WASH committees, to form larger cooperatives and register with local and district authorities. By linking these local-level institutions to higher-level authorities, the project sought to give communities better access to information and resources from district and national governments.

The project also supported the construction of facilities and small-scale structures within communities to enhance a number of different aspects of resilience. These efforts were typically coordinated with existing government programmes including, for example, the Open Defecation Free (ODF) initiative and the Poverty Alleviation Fund (PAF). The project helped build WASH facilities in its working communities, including the construction of wells, deep boreholes, and toilets. Also, to address disaster risk-reduction, flood defences and evacuation routes were improved by building embankments, constructing culverts, and planting trees.

A number of the project activities were aimed at building the capacity of VDC-, district-, and national-level governments to respond to disasters and improve WASH outcomes themselves. As described above, this was, in part, a case of creating more formal linkages between existing community institutions and higher-level authorities. However, the project also provided training and organised exchange visits to increase the capacity of VDC and district government personnel, placing special emphasis on creating disaster risk-management plans (DRMPs) and WASH plans. As part of this planning process, VDC- and district-level authorities were encouraged to set aside funds from their budgets for disaster risk reduction activities and emergency response.

The project also dedicated some funds to the training of partner and Oxfam staff. These activities focused on improving plans for humanitarian crises, making cash transfers during crises, and building expertise around topics such as Gender in Emergency (GiE) – to mainstream gender in emergency situations and ensure minimum standards are achieved.

Finally, the project sought to influence government policies at both the national level and lower levels. One particular goal of these advocacy activities was to enable the national government to pass the Disaster Management Bill, which, in spite of numerous attempts, had failed to go through the Nepalese parliament. The project provided technical and financial support to the National Network of Community Disaster Management Committees (NCDMC), giving special guidance on advocacy strategies.

2.2 PROJECT LOGIC AND INTENDED OUTCOMES

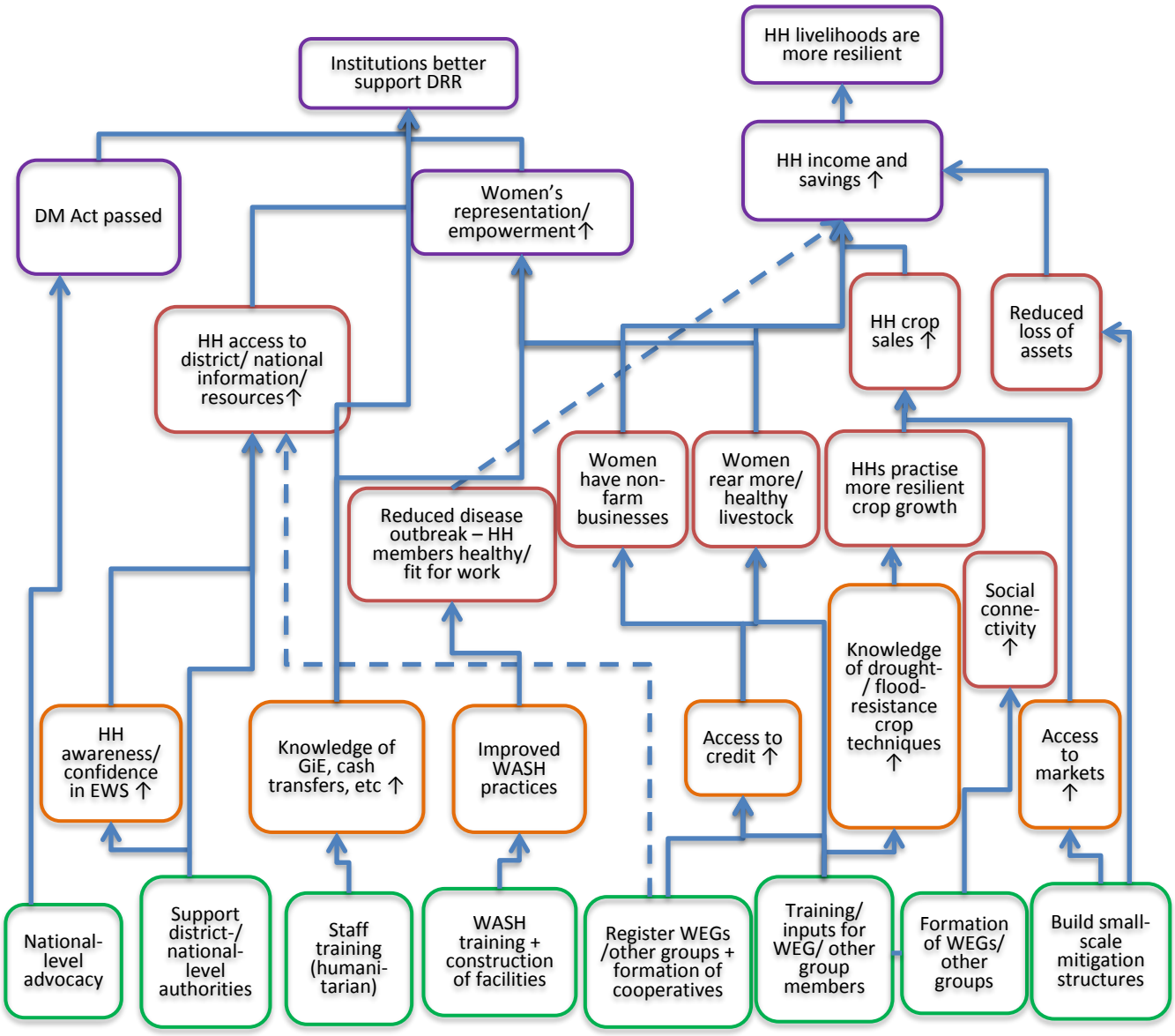
In this section, we describe how the project was supposed to achieve its goals. Using existing documentation about the project as well as discussions with the team implementing the project, we can map out the intended causal links from project activities (green), via outputs (orange) and intermediate outcomes (red), to overall resilience outcomes (purple). This results in the Logic Model shown in Figure 2.1. It should be noted that this diagram stops at the factors that could be considered drivers or characteristics of resilience, and does not include the final outcomes beyond resilience (such as improved well-being and realisation of rights), that the project may have been trying to promote in the long run.

The formation of and support to WEGs was intended to build resilience through a number of different channels. Firstly, the trainings provided were varied, partly focusing on building the resilience of more traditional livelihood strategies – such as growing crops and rearing livestock – but also promoting non-farm businesses. It was hoped that household members would build up a base of knowledge and understanding around resilient livelihood activities and then apply this in practice. It was also expected that, by building up the savings and credit functions of the WEGs, households would have better access to credit and therefore be better able to try new income-generating activities, especially if WEGs were better linked to each other though cooperatives. These extra livelihood strategies were intended not only to boost household income and savings, but also to give women more control over household resources, in turn empowering women and improving their representation. It was further anticipated that the formation of the WEGs, and indeed other groups like the CDMCs, would help build social connectivity within and between project communities.

The construction of small-scale flood mitigation structures sought to build more resilient sources of household income through two channels. Firstly, these structures were designed to directly protect crops and assets from flood water. Secondly, however, it was also hoped that these structures would maintain access to markets by protecting roads and tracks when floodwaters are high.

Another intended consequence of formalising and registering WEGs and other community groups was to give project households better access to information and resources from the district and national governments, especially during shocks and stresses. Linking up CDMCs with the equivalent institutions at the VDC, district, and national level was seen as especially important as a means of ensuring that communities' plans were well integrated and that early warning information could be collected and disseminated effectively.

Figure 2.1 – Logic Model for the Joint Programme



Key: Project Activities; Outputs; Intermediate Outcomes; Resilience Outcomes

In some sense the project's WASH activities affect both crisis and non-crisis periods, broadening the scope of the project beyond what would typically be understood in terms of disaster risk reduction and livelihoods aspects of resilience. However, the WASH elements of the project were intended, in particular, to reduce the outbreaks of disease during floods by improving households' WASH practices. Insofar as having healthy and fit household members depends on the disease environment of the community at large, this also contributes to the livelihoods outcomes of the project.

Finally, support provided to district- and national-level authorities, as well as the national-level advocacy work, sought not only to improve directly the disaster risk reduction capacity of institutions at different scales, but also to create changes at the household level. With government authorities focusing more on disaster risk reduction, households in project communities were supposed to have better access to back-up resources and early warning information. This, in turn, was expected to increase households' awareness of and confidence in early warning systems and the disaster preparedness plans put together by district and national governments.

The activities pertaining to the training of Oxfam and partner staff are included in Figure 2.1 for completeness.

3 EVALUATION DESIGN

The central problem in evaluating the impact of any project or programme is how to compare the outcomes that resulted from that project with *what would have been the case* had the project or programme not been carried out. In the case of this Effectiveness Review, information about the situation of households in the project communities was collected through a household questionnaire, but clearly it was not possible to know what their situation would have been had the project activities not been undertaken. In any evaluation, this 'counterfactual' situation cannot be directly observed: it can only be estimated.

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

This evaluation focuses on assessing both household- and VDC-level impacts of the project. Therefore, we aim to compare the direct beneficiaries within project communities with similar households in similar non-project VDCs.

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

However, in the case of the project examined in this Effectiveness Review, neither the VDCs where the project was implemented, nor the participant households within those VDCs were selected at random.

The implementers targeted VDCs that were poor and were more vulnerable to shocks and stresses. In part, this assessment of VDCs' vulnerability related to geographical factors, such as proximity to rivers, which would directly determine whether or not community members were under threat from flood. Other anthropocentric factors were also taken into account. For example, VDCs' dominant livelihood strategies would clearly determine vulnerability to droughts, cold waves, and disease, while population density would determine the extent to which households were threatened by fires.

However, within the same districts, there were a number of other VDCs with similar characteristics that faced similar risks, but which were not included in the Joint Programme. This allowed a 'quasi-experimental' evaluation approach to be adopted, in which the situations of households in VDCs not included in the project – in so-called 'comparison' sites – were assumed to provide a reasonable estimate for the counterfactual of households who participated in the project.

It is important to note that within the project VDCs, those who participated in the household-level activities of the project were not a random cross-section of residents. In this evaluation, we elected to focus on the participants in the WEGs. Not only were a large proportion of the project benefits channelled through these groups, but also the households participating in the WEGs tended to participate in the CDMCs as well. Vitally, the project partners had maintained lists of WEG members that could be used to track project participant households. Finally, households were selected into the

WEGs according to clear criteria. In particular, the project targeted poor households with low levels of education.

In fact, since the populations within the targeted VDCs were reasonably homogeneous in terms of wealth and education levels, the majority of households were eligible to participate in the WEGs. However, only around 20 percent of households contained individuals that were members of WEGs. This implies that there may have been factors other than wealth and education levels determining which households participated in WEGs.

As such, it is likely that the households who participated in the activities that were channelled through the WEGs differed from their non-participant neighbours – for example, in terms of their wealth, household composition, their sense of initiative, willingness to take risks, or in their social connections. It was therefore necessary to try and identify similar households within the comparison communities to create a suitable counterfactual for the project participants.

We attempted to resolve this issue in two ways. Firstly, we randomly sampled from the population in comparison VDCs, but then asked 'screening questions' at the start of the questionnaire to ensure households were not too wealthy or too highly educated to have been eligible to participate in the WEGs in project VDCs. However, the success of this method relies on knowing all of the criteria that determine whether or not households would have participated in WEGs, had the project operated in their VDC. In the event, the project's eligibility criteria did not entirely determine which households participated in the project, so there may have been further differences between project participants and their neighbours, which are not accounted for by using screening questions alone. We explain this process in more detail in Section 4.

Secondly, the statistical analysis used in this evaluation also, allowed us to improve the confidence in our comparison between those households that did and did not participate in the project. This is especially important given the limitations in using screening questions described above. Households in the project communities were 'matched' with households with similar characteristics in the comparison communities. Matching was performed on the basis of a variety of observable characteristics – including household size, education levels, productive activities, and indicators of material well-being, such as housing conditions and ownership of assets. Since some of these characteristics may have been affected by the project itself (particularly those relating to productive activities and wealth indicators), matching was performed on the basis of these indicators *before* the implementation of the project. Although baseline data were not available, survey respondents were asked to recall some basic information about their household's situation in 2010, before the project was implemented. This recalled baseline data is unlikely to be highly accurate. However, it still serves as a suitable proxy for households' baseline situation, enhancing the reliability of the comparisons made in this report.

The survey data provided a large number of baseline household characteristics on which matching could be carried out. (The characteristics that were in fact used are listed in Appendix 3.) In practice, it is very difficult to find households in the comparison communities that correspond exactly in all these characteristics to households in the project communities. Instead, these characteristics were used to calculate a 'propensity score' – the conditional probability of the household participating in the project, given particular background variables or observable characteristics. Households in the project and comparison communities were then matched based on this propensity score. After matching, it was possible to test whether the distributions of each baseline characteristic were similar between the two groups. Technical details on this approach are described in Appendix 3.

As a check on the results derived from the propensity-score matching process, results were also estimated using multivariate regression models. Like propensity-score matching, multivariate regression also controls for measured differences between the 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. The regression models tested are described in Appendix 4.

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 – such as individuals' attitudes or motivation, differences in local leadership, weather, or other contextual conditions – then estimates of outcomes derived from them may be misleading. This is a cause for particular caution when in evaluating a project in which participants were to some extent self-selected. This point is further discussed in Sections 5 and 6, when interpreting the statistical results.

4 DATA

4.1 RESPONDENTS INTERVIEWED

To form a sample of project and comparison households for this evaluation, we began by selecting the project communities on which to focus the evaluation. All four project districts were included in the sample, ensuring that the work of all the main project partners was included in this evaluation.

Within each district, two VDCs were selected for the survey.[1] Several criteria were used to determine which VDCs were suitable for the evaluation. Overall, we targeted VDCs where the project had been implemented to the fullest extent, that is, the complete range of household- and community-level activities had been undertaken. Thus, our evaluation considers how project participants are affected by all the project work taken together, but we cannot easily disaggregate these effects by each activity. For example, we cannot examine the impact of WASH training independently of the other project activities.

There were also some practical constraints, which prevented fieldwork from being undertaken in certain project areas. Accessibility to a number of VDCs was limited as the survey was carried out during a period of high ethnic tensions with widespread protests and strikes, associated with the new Constitution of Nepal coming into effect late in 2015. Collecting data in these areas would have put enumerators and supervisors at an unacceptable level of risk.

In order to establish which VDCs would be suitable for comparison purposes, a list of key socio-economic and geographical characteristics was drawn up, which was based on the original criteria used to determine whether VDCs could participate in the project. These characteristics were mainly related to the VDCs' vulnerability to risks, including:

- the dominant livelihood strategies employed by VDC residents
- the distance of the VDC from large rivers
- ethnic composition of VDC members
- overall VDC wealth levels
- level of out-migration from VDCs (e.g. to India, in search of work).

Through discussions with project staff, we were able to identify a total of 12 suitable comparison VDCs, three in each district.

During these discussions, we also mapped out the other government and NGO initiatives that were operating in the project area. There were two key government projects that worked in the majority of VDCs within the project districts. The Open Defecation Free (ODF) initiative seeks to improve WASH outcomes in line with the sanitation Millennium Development Goal, mainly by working with local groups and institutions to construct toilets and train community members on hygiene practices. In addition, the Poverty Alleviation Fund provides an array of interventions aimed at social mobilisation and strengthening livelihoods among marginalised populations in Nepal.

If the comparison households in our sample were participating in the ODF and PAF initiatives, but project households were not, and this was not taken into account in our analysis, the estimated impact of the project would be biased downwards. Conversely, if comparison households were not participating in the ODF and PAF initiatives, but project households were, then we may risk overestimating the project's impact. We actually sought to only *include* VDCs in our sample where the ODF and PAF initiatives

were operating. As a result, the presence of government activities is consistent between the project households and the comparison households. Thus, by comparing these two groups, we are still able to estimate the effect of the project, over and above the work undertaken by the ODF and PAF programmes. According to project and partner staff, there were no other major NGOs operating in the project and comparison VDCs selected for our sample.

Within the project communities, the current lists of individuals involved in WEGs were used for sampling. In particular, the field supervisor met with the relevant project partner and local leaders in each VDC to convert the list of individual names into a condensed list of households. In each of the intervention VDCs, 35 households were targeted for interview.[2] These were selected by applying systematic random sampling to the full list of households.

A similar procedure was applied in the comparison VDCs. In order to create lists of households for each VDC, the survey team used lists of voters provided by the Electoral Commission of Nepal. Again, these lists of individual voters were compressed to a list of households through discussion with local leaders. In the comparison VDCs, 44 households were targeted for interview, and these were selected using systematic random sampling.

To help ensure that the comparison households had similar characteristics to the project households, we used the same criteria that were used to select households for the project to screen out unsuitable households from the comparison VDCs. In particular, we sought to exclude comparison households that were both too wealthy and too educated to have been eligible for participation in the project. We asked households two direct 'screening questions' before starting the main survey, which proxied for the project's initial eligibility criteria, to determine whether or not they would suitable for the sample:

1. "Does this household own more than 20 katthas of land?"
2. "Does every adult member of this household have more than four years of formal education?"

Any households in the comparison VDCs that met *both* of these conditions, were excluded from the sample. These screening questions were discussed with the project staff and piloted in the field test of the questionnaire.

As we noted in Section 3, these criteria did not exclude a large proportion of the households from the project VDCs because the VDCs' populations were fairly homogenous in terms of wealth and education levels. The majority of households were poor and had low levels of education. However, only 20 percent of households in the project VDCs participated in WEGs. This means that there may have been other factors, besides the eligibility criteria outlined above, that determined which households did and did not participate in the WEGs. As such, it is important to check and control for baseline differences between the project and non-project households in our sample, as we discuss in Section 4.2.

Given the rigidity of the sampling strategy described above, not all sampled households could be reached for interview straight away. If no one from a particular sampled household could be located, the enumerator was asked to make one revisit. If this revisit was unsuccessful, the survey team could draw a replacement household. In the project areas, this was the simply the next household on the lists that were created before the survey work began. In the comparison areas, enumerators were asked to re-sample the next closest household, in terms of distance. This replacement procedure was also used if a comparison household was excluded from the sample according to the screening questions described above.[3]

The numbers of households interviewed in the project communities and in the comparison communities are shown in Tables 4.1 and 4.2 respectively.

Table 4.1: Intervention communities and numbers of households interviewed

District	Community	Households interviewed
Dhanusha	Itaharwa	33
	Jhatiyahi	36
Rautahat	Banjaraha	35
	Katahariya	36
Salarhi	Attrauli	34
	Kalinjor	35
Saptari	Launiya	35
	Ramapuramalhaniya	36
Total		**280**

Table 4.2: Comparison communities and numbers of households interviewed

District	Community	Households interviewed
Dhanusha	Lakhouri	43
	Mithileswormauwahi	45
	Paudeswor	44
Rautahat	Badharwa	43
	Bariya	44
	KarkachKamaiya	47
Salarhi	Dhurkauli	45
	Parwanipur	42
	Sasapur	45
Saptari	Bishariya	43
	Joginya 1	45
	Ko.Madhepura	43
Total		**529**

The data for this evaluation were principally collected at the household level. Questionnaires were conducted with a particular household member, but they were asked to answer questions for the household as a whole. A household was defined as those individuals who normally (in the last three months) shared the same cooking facilities.

To gain the most accurate picture of the household's situation, the survey team targeted the household head for interview. The household head was defined as the person who mainly makes decisions on the education of children, on household finances, and on household expenses. This definition was developed through consultation with partner staff to ensure it was appropriate to the local context. If the household head was not available for interview, the next most senior adult was interviewed.

Before the survey started, respondents were given some basic information about the purpose of the survey, to help manage their expectations. The enumerators explained

that the survey was being undertaken to help Oxfam better understand the lives of people in the community, and that it was for 'research purposes only'. It was also made clear that no special support would come to households as a result of the answers to questions in the survey.

Interviews were carried out using mobile devices. The questionnaire was created in a piece of Open Data Kit software, called *SurveyCTO*, and then downloaded onto a mobile phone given to each enumerator. The functionality of the mobile phones was reduced so that they could only be used for data-collection purposes. The data were uploaded nightly by field supervisors and checked by the evaluation team to ensure high data quality.

4.2 ANALYSIS

Before analysing the effects of the project on resilience outcomes, we compared project participant households and comparison households in terms of their demographic characteristics, livelihoods activities and economic situation in 2010 (that is, before the Joint Programme began). This helped to check the suitability of the comparison group, and ascertain what variables could be included in the main analysis to control for observable differences between project and non-project households.

Some of the data were based on information recalled during the questionnaire. Before beginning the main part of the questionnaire, the enumerators worked with the respondents to establish a suitable event or season from 2010, to help them to think consistently about the correct time period when answering the recall questions. However, given the difficulties of remembering specific aspects of livelihoods, asset ownership, and other activities, it is possible that these recall questions may be subject to measurement error. This should be borne in mind throughout the analysis.

The full comparison of project participant households and comparison households in terms of all these characteristics is shown in Appendix 2. There are four key differences that we wish to highlight.

Firstly, literacy and education levels appear to be somewhat higher among the project households than the comparison households. In the comparison group, just 29 percent of the households have a household head that can read and write a simple letter, compared with 46 percent for the project households. In part, this may be because the selection of less-educated households was applied more strictly by the screening questions in our survey than by the project.[4] However, it seems this would be unlikely to account for the full difference as the screen we applied was only designed to exclude households where education levels were high across *all* adult members.

Secondly, there are clear ethnic differences between project and non-project households. Most strikingly, the proportion of Dalit households is approximately 46 percent in the comparison VDCs, compared to just 27 percent in the project VDCs. This is important because ethnicity may capture cultural norms, preferences over livelihood strategies, and aspects of social connectivity that would otherwise not be observed in our survey. It is somewhat surprising these differences exist, given that ethnic composition was explicitly taken into account when selecting comparison VDCs.

Thirdly, the income sources on which households relied in 2010 appeared to differ between project and comparison households. While 73 percent of project households engaged in farming crops in 2010, the figure was just 50 percent for the non-project households. However, casual labour – such as construction, carpentry, and masonry – was far more prevalent among the comparison households than the project households. Interestingly, these differences in livelihood strategies do not appear to

correspond to significant imbalances in the wealth distribution between project and non-project households.

Finally, our data suggest that participation in community groups was substantially more prevalent among the project VDCs than non-project VDCs. Around 55 percent of the project households participated in some type of community group in 2010, compared to just 25 percent in the comparison group. These differences must be carefully considered, given the importance of community groups to the project's activities, and therefore to the conclusions we can draw from this evaluation.

Any differences between project and comparison households that existed before the project have the potential to bias comparisons of the project's outcomes between the members of the project and comparison respondents. Therefore, we tried to control for these baseline and demographic differences when making such comparisons. This was especially important for livelihoods and wealth, which could be regarded as potential outcomes of the project – we hope to find out whether the project affected these outcomes, rather than there simply being differences between the project participants and the comparison group.

Some of the differences between the project participant households and the comparison group identified above may be down to recall error. However, this would require the project participants to systematically overstate their livelihood diversification and wealth and/or the comparison households to systematically understate. There are several possible reasons this might have happened in our survey. In particular, it may have been difficult for project respondents to remember back to a time *before* any project activities, so their recall answers may include some mix of their baseline status and the effects of the project. This is especially likely for questions pertaining to group participation because it may be hard for respondents to remember precisely when a certain community group was set up, especially if scoping activities occurred before the 'true' start of the project.

However, in the absence of these types of *systematic* biases for the project participant households and comparison households, any measurement error that arises due to recall would actually lead any differences between project and non-project to be *underestimated*.[5] Thus, it is unlikely that there are truly no differences between project households and comparison group, in terms of literacy, ethnicity, livelihood strategies, and group participation (in 2010). As such, we believe it is important to control for these differences in our analysis.

As described in Section 3, the main approach used in this Effectiveness Review to control for the baseline differences was propensity-score matching (PSM). The variables on which respondents were matched were selected from among the full list detailed in Appendix 2, based on two key factors. Firstly, we selected those variables that were thought to be the most significant in influencing respondents' participation in the project. Secondly, we aimed to include variables that could affect potential project outcomes *as well as* the likelihood of participating in the project.

In particular, the literacy of the household head, as well as their ethnicity, was used for matching, as were variables to capture households' engagement in livelihood activities and participation in groups in 2010. This was in spite of concerns about the degree of recall error in those data. The list of matching variables selected and the full details of the matching procedure are described in Appendix 3.

After matching, project participant households and comparison households appeared to be reasonably well balanced in terms of each of the selected variables. One caveat is that 8 of the 280 project participant households in the sample and 2 of the 439 comparison households could not be matched and had to be dropped from the analysis. Consequently, the estimates of the project's impact presented in Section 6

are not based on the whole population interviewed, but exclude a non-random minority. However, the exclusion of only 10 observations at this stage of the analysis is unlikely to make a substantial difference to our main findings.

We also check whether the different sampling procedures used for the project and non-project VDCs may have led to imbalance in the specific household characteristics – land and education – that were used to select project participants and screen out households that were too wealthy or too educated from the comparison group. However, as shown in Appendix 3, the matched sample appears to be balanced in terms of land holdings in 2010, which was measured using recalled data, and the education levels of adult household members. This should reassure us that, even if the screening process for the comparison group does not perfectly replicate the procedure that was originally used to select project participants, this does not bias our results for the matched sample.

All the results described in Section 6 of the report were tested for robustness by estimating them with several alternative statistical models, including alternative PSM models and linear or probit regression models. These robustness checks are shown in Appendix 4. However, the results of the alternative PSM and regression models generally produced estimates of outcomes that were similar in magnitude and in statistical significance to those derived from the original PSM model. The few cases where the models produced divergent results are discussed in Section 6, in the text or in endnotes.

As mentioned in Section 3, PSM and regression models can only control for the baseline differences between project and comparison households for which data was collected in the survey. If there are any ‘unobserved’ differences between the two groups – such as individuals’ attitudes or motivation, differences in local leadership, or weather, or other contextual conditions – then these may bias the estimates of outcomes described in Section 6. The evaluation design and the selection of respondents were intended to minimise any potential for unobserved differences, but this possibility cannot be excluded, and must be borne in mind when interpreting the results.

5 MEASURING RESILIENCE IN NEPAL

The Joint Programme was specifically aimed at increasing households' resilience. As part of its Global Performance Framework, Oxfam GB has developed an innovative approach to measure the resilience of households to shocks and stress and their ability to adapt to change.[6] This approach involves capturing data on various household and community characteristics falling under the five interrelated dimensions presented in Figure 5.1.

Oxfam defines resilience as 'the ability of women and men to realise their rights and improve their well-being in spite of shocks stresses and uncertainty'. One reason why measuring resilience is challenging is that we can only really assess whether a system has successfully coped or adapted after the fact.

In this Effectiveness Review, we were partially able to observe how well households had coped with shocks, stresses, and uncertainty, because the survey work was carried out during a period of drought for the entire region. Certain households also experienced cold waves, fire, and localised flooding in the months preceding the survey. In Section 6, we consider how well households maintained their livestock and crops, and their level of household wealth, in spite of these shock and stresses.

However, looking at these sorts of final outcomes is not sufficient to tell us about the project's full impact on resilience for two main reasons. Firstly, only a small proportion of households was subject to shocks other than drought. In particular, just 13 percent of households reported experiencing a flash flood in the two years prior to the survey. Since many of the project activities sought to build resilience to flooding, looking at final outcomes alone would give a very incomplete picture of the project's impact. Moreover, focusing on previous shocks is backward-looking, and does not allow us to investigate the project's impact on resilience in the *future*.

The characteristics approach to resilience measurement, which we adopt in this section, is based on the assumption that there are particular characteristics of households and communities that affect how well they are able to cope with shocks and positively adapt to change. Insofar as there are multiple final well-being outcomes, about which we are ultimately concerned, there should also be a wide range of resilience characteristics. Where possible, we also wish to conceptualise resilience as operating at many different levels (individual, household, community, and so on) as well as for different shocks with different time horizons. As a consequence, the number of resilience characteristics is potentially very high. A limitation, of course, is that we do not know for certain how relevant particular characteristics actually are; rather, we assume they are important based on common sense, theory, and an understanding of the local context.

To help structure discussions around the list of characteristics that inform the overall measure of resilience, we have typically used five 'dimensions'. These are presented in Figure 5.1.[7]

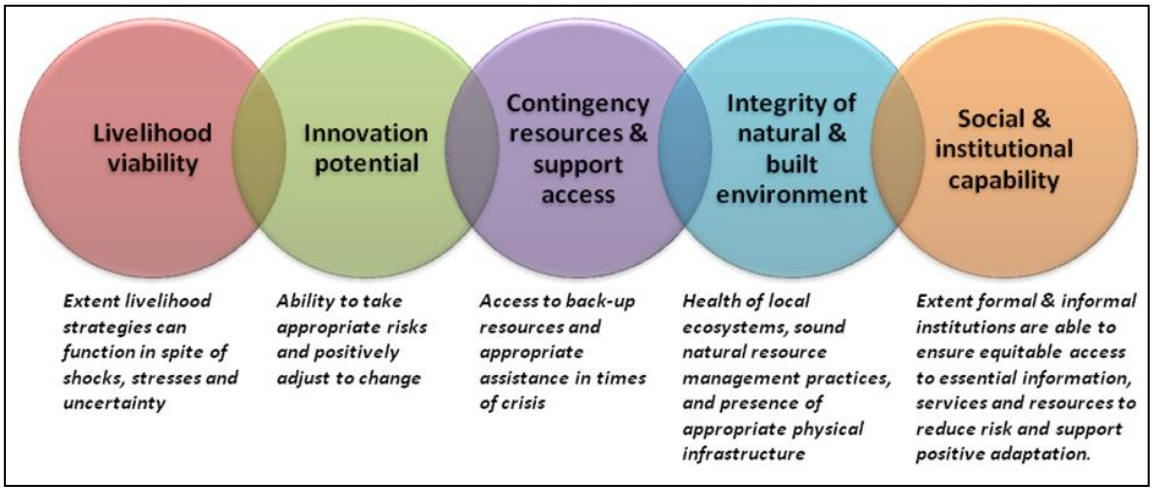
First, if we think about what a household would need in order to cope with current and future shocks, stresses and uncertainty, a **viable livelihood** is likely to be one of them. If a shock happens, for instance, a household dependent on just one precarious livelihood activity is likely to be more negatively affected than another that has one or more less sensitive alternatives to fall back on, *all other things being equal*. In addition, households that are on the margins of survival are less likely to be resilient than their

relatively more wealthy counterparts. Where longer-term climatic trend prediction information exists, it is also important to assess how viable current livelihood strategies would be given the range of likely future climatic scenarios.

Innovation potential focuses on a household's ability to positively adjust to change, whether anticipated or not. We can hypothesise that such potential is dependent on factors such as the knowledge and attitudes of relevant household members themselves, their ability to take risks, and their access to weather prediction, market information and relevant technology and resources.

Moreover, there will likely be times when even households with the most 'resilient' and adaptive livelihood strategies will find it tough to get by. **Access to contingency resources and external support** – e.g. savings, food and seed reserves, social protection, kin and non-kin support networks, and emergency services – are, therefore, likely to be critical in supporting households in coping with shocks and positively adjusting to change.

Figure 5.1: Dimensions affecting the ability of households and communities to minimise risks from shocks and adapt to emerging trends and uncertainty



It is further recognised that **healthy ecosystems** are better able to cope and adjust to climatic shocks/change than those that are relatively more degraded. We may reasonably assume – again, with all other things being equal – that households whose livelihoods are dependent on healthier ecosystems will be in a better position to adjust to climatic shocks/change than those that are not. The presence of appropriate infrastructure (e.g. pit latrines and roads) that is resilient to shocks and stresses (e.g. flooding) is equally important; if critical infrastructure no longer functions, or collapses in times of shocks and stress, the livelihoods and/or health of community members can be negatively affected.

In most, if not all cases, it is necessary to look beyond the household level when examining resilience and adaptive capacity. Indeed, it is reasonable to assume that households are likely to be better able to successfully adjust to climatic shocks/change when they are part of larger coordinated efforts at the community level and beyond. The **social and institutional capability** dimension, in particular, is concerned with the effectiveness of informal and formal institutions in reducing risk, supporting positive adaptation, and ensuring equitable access to essential services in times of shock/stress. In the absence of this capability, we can assume that community-level duty bearers will be less effective in fulfilling their responsibilities in supporting community members to reduce risk and/or successfully adapt.

While the five dimensions of resilience described here provide an overall framework, the challenge in creating a measure of resilience is to identify specific characteristics

that are appropriate to the local context. For this Effectiveness Review, we consulted local staff from each of the different partner organisations and from Oxfam to identify what factors they considered the most important for contributing towards resilience within the project area. Two focus groups were then carried out, one entirely comprised of women and one entirely comprised of men, in a community that was similar to those included in the survey.[8] This enabled us to further probe people's understanding of what factors contributed to their resilience.

This process led to a set of characteristics of resilience being identified, listed in Table 5.1. It is important to note at this stage that while not all characteristics considered in this Effectiveness Review may be directly linked to the project activities, all are deemed to be important to a household's overall resilience in this particular context. The right-hand column of Table 5.1 shows the characteristics on which the project was expected to have an impact, given its Logic Model.

Table 5.1: Characteristics of resilience examined in this Effectiveness Review

Dimension	Characteristic	Connected to project logic?
Livelihood viability	Ownership of productive assets	Yes
	Ownership of land	No
	Ownership of livestock	Yes
	Livelihood diversification	Yes
	Crop diversification	Yes
	Cultivation of drought-resistant crops	Yes
	Dietary diversity.	No
	WASH attitudes	Yes
	Livestock hygiene practices	No
Innovation potential	Attitude to change	No
	Access to credit	Yes
	Awareness of climate change	No
	Adoption of innovative practices	No
	Access to markets (in monsoon season)	Yes
Access to contingency resources and support access	Participation in community groups	Yes
	Social connectivity	Yes
	Savings	Yes
	Formal earnings	No
	Awareness of community disaster management plan	Yes
	Awareness of VDC disaster management plan	Yes
	Awareness of district management plan	Yes
	Direct access to emergency items	Yes
Integrity of natural and built environment	Availability of drinking water	Yes
	Irrigation for agriculture	Yes
	Protection of house from flood	No
	Protection of house from fire	No
	Use of organic fertiliser	Yes
	Caring for forest/greenery	No
Social and institutional capability	Participation in disaster planning process	Yes
	Early-warning system	Yes
	Awareness of local leaders' planning activities	Yes
	Women's representation in the community	Yes
	Registration of community groups	Yes

The questionnaire used in the Effectiveness Review included questions relating to each of the characteristics listed in Table 5.1.

Data from these various indicators of resilience were aggregated using an approach similar to the Alkire-Foster method, adapted from the methodology used by the Oxford Poverty and Human Development Institute for measuring multidimensional constructs, such as poverty and women's empowerment. For each characteristic, a benchmark was defined based on what it means for a household to be faring reasonably well in relation to the characteristic in question. The particular benchmarks used for each characteristic are detailed in Appendix 1. For example, each household was defined as scoring positively in terms dietary diversity, if household members consumed a carbohydrate source on each of the seven days preceding the survey, a protein source on at least three of those days, and vegetables on at least three of those days. These cut-offs were developed through conversations with project staff and by checking the summary statistics for each variable to ensure the proposed thresholds were not obscuring important variation in the data. There is, however, inevitably a degree of arbitrariness in defining such cut-offs. Alternative cut-offs and formulations of the indicators were tested as a check on the robustness of the results obtained from applying the cut-offs.

Having used the cut-offs to create a binary variable for each characteristic of resilience, it was then necessary to find some way of aggregating across all the indicators of resilience. One of the biggest challenges in constructing this type of index is finding ways to *weight* different indicators – or indeed different dimensions – against one another to construct an index that adequately reflects what is important for resilience in the local context. We adopt four different strategies for assigning weights to the indicators in this Effectiveness Review, to ensure that our findings are not sensitive to different assumptions about what truly matters for resilience.

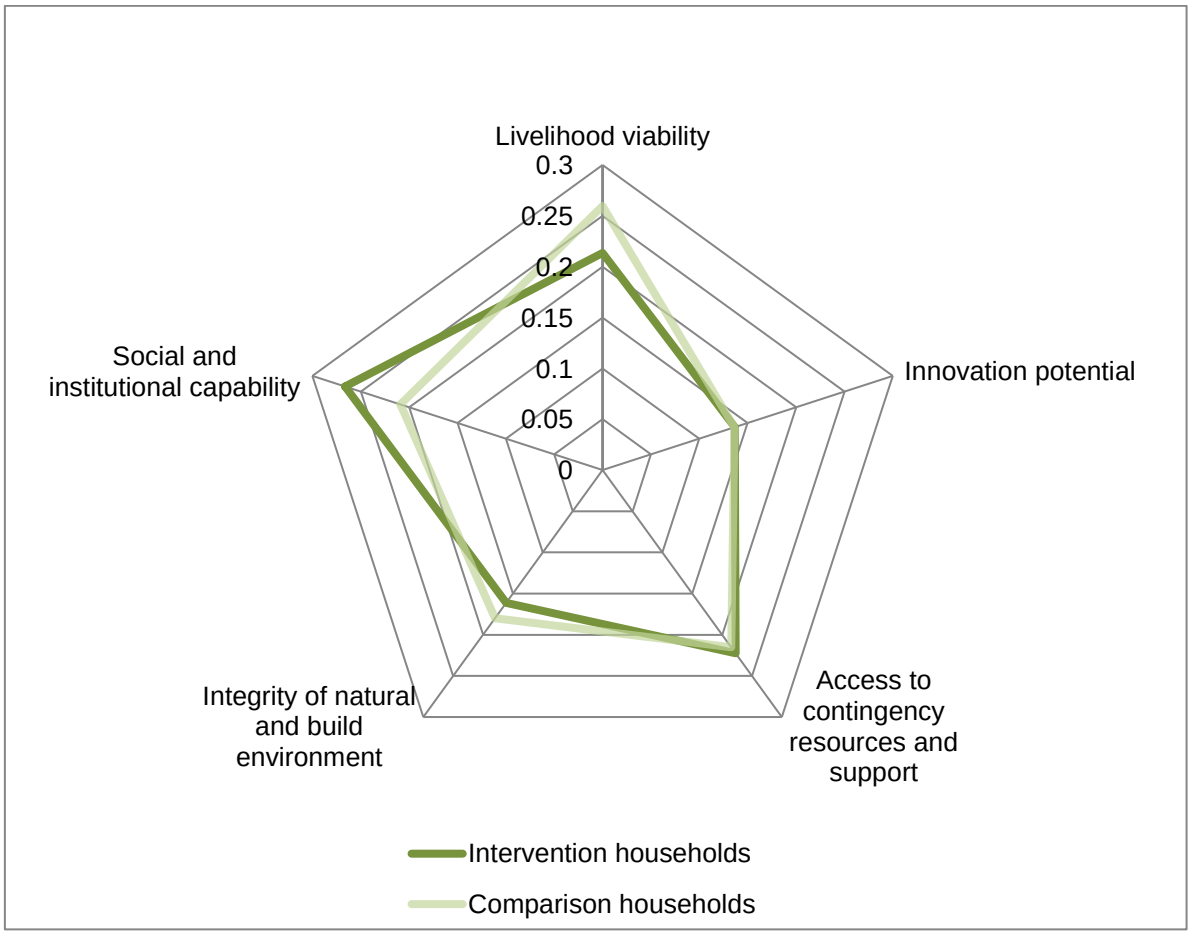
Firstly, in line with previous Effectiveness Reviews, we weight the indicators equally, by simply counting the proportion of characteristics in which the household scored positively. We refer to this measure as the *base resilience index*.

Secondly, we create an index, which is based on the idea of giving equal weight to each of the five resilience dimensions. To do this we first created a separate index for each dimension by counting the proportion of indicators under that dimension in which the household scored positively. Then we took the average across these dimension-specific indices to give an overall resilience index. We refer to this as the *equal dimensions resilience index*.[9]

The third and fourth methods we use to weight the resilience indicators utilise a special module that was included in the questionnaire to try and extract households' preferences over what they thought was most important for resilience. This involved showing respondents a special card, on which pictures had been drawn to represent each of the five dimensions. This tool was adapted from a previous Effectiveness Review through consultation with the survey team, to ensure it suited the local context, and is shown presented in more detail in Appendix 6. Enumerators were trained to explain each picture carefully. Respondents were then given 15 stones, which they could allocate to each picture in accordance with their perceived importance of each resilience characteristic. Specifically, they were asked: 'How important are the following things to you in making sure your household member have everything they need, even in difficult times?'

The average weightings that respondents placed on each dimension are shown in Figure 5.2, separating out the intervention and comparison groups, and using the matched sample. We divide the number of stones placed on each dimension by 15, so that the weights shown in Figure 5.2 add up to one.

Figure 5.2: Results from the weighting exercise



Across both the project and non-project households in the sample, livelihood viability, access to contingency resources and support, and social and institutional capability are given more weight by respondents than the other two dimensions. There are, however, some differences between the intervention and comparison groups:

1. Project households weight livelihood viability *less* than non-project households. The average weight assigned to this dimension was 0.21 for the intervention group, and 0.26 for the comparison group.
2. Project households weight integrity of the natural and built environment *less* than non-project households. The average weight assigned to this dimension was 0.16 for the intervention group, and 0.18 for the comparison group.
3. Project households weight social and institutional capability more than non-project households. The average weight assigned to this dimension was 0.27 for the intervention group, and 0.21 for the comparison group.

These results therefore imply that the project was able to influence respondents' deep preferences, changing what they believed was important for their resilience. The extra emphasis that project households place on social and institutional capability is particularly striking because the project channelled so many of its activities through community groups and local disaster risk management committees. It appears that this mainly came at the expense of the weighting given to livelihood viability.

Using this data, we created two further resilience indices. The *personal dimension resilience index* weights each dimension according to that particular respondent's answers to the weighting exercise. We may, however, believe that assigning different dimension weights for each household affects the comparability of our data, especially given the differences between intervention and comparison group described above. To overcome this concern, we also show the *sample dimension resilience index*, which uses the average values of the dimension weights for the sample as a whole.

The project's effect on resilience, as measured through these indices, is reported in Section 6.9.

6 RESULTS

This report is intended to be free from excessive technical jargon, with more detailed technical information being reserved for the appendixes and endnotes. However, there are some statistical concepts that cannot be avoided when discussing the results. In this report, results will usually be stated as the average difference between the project households (that is referred to as the 'intervention group') and the matched non-project households (named the 'comparison group'). In the tables of results on the following pages, statistical significance will be indicated with asterisks, with three asterisks (***) indicating a *p*-value of less than 1 percent, two asterisks (**) indicating a *p*-value of less than 5 percent and one asterisk (*) indicating a *p*-value of less than 10 percent. The higher the *p*-value, the less confident we are that the measured estimate reflects the true impact, as opposed to simply random variation in the data. Results with a *p*-value of more than 10 percent are not considered to be statistically significant.

6.1 INTRODUCTION

This section presents a comparison of the households interviewed in project and comparison communities in terms of various outcome measures relating to the project under review. As described above, asterisks are used in the results tables to indicate where the differences are statistically significant at at least the 10 percent significance level.

The results are shown after correcting for observed baseline differences between the households interviewed in the project communities and those in the comparison communities using a propensity-score matching (PSM) procedure. This means that when we report differences in the means for the intervention group outcomes and comparison group outcomes, *this is for the matched sample*. More information about the procedure applied is found in Appendix 3. All outcomes discussed here have also been tested for robustness with alternative statistical models, as described in Appendix 4. Where those alternative models produce markedly different results from those shown in the tables in this section, this is discussed in the text or in endnotes.

There are two key limitations to our analysis, which have been described above, but are repeated here because they affect the interpretation of our results:

1. A non-random minority of households were excluded from the analysis during the matching process (8 of 280 intervention group households and 2 of the 529 comparison group households). This means that the results shown in the tables in this section are not based on a fully representative sample of households in the project communities.[10]
2. There may be 'non-observable' differences between the project participants and comparison households – such as individuals' attitudes or motivation, differences in local leadership, weather or other contextual conditions. If these unobserved differences also influence the potential outcomes we consider in this section, then our estimates of the projects' effects will be biased. This possibility must be borne in mind when interpreting the results.

6.2 INVOLVEMENT IN PROJECT ACTIVITIES

The first step in understanding the project's impact is to examine the extent to which respondents reported having received the types of support and having participated in the various activities implemented by the project.

In Table 6.1, we show the differences between the project participant households and the comparison households in terms of their participation in community groups. Respondents were asked whether they or any member of their household attended meetings of a series of community groups at the time of the survey. We first show an indicator for whether or not the household participates in *any* community groups in Column 1, followed by the number of community groups in which the household participated in Column 2. The remaining columns report whether or not household members participate in specific types of community groups.

Table 6.1: Participation in community groups

Part A

	1	2	3	4
	Household member(s) participate in any community groups	Number of community groups in which household participates	Household member(s) participate in WEG	House member(s) participate in other women's group
Intervention group mean	0.88	4.41	0.74	0.57
Comparison group mean	0.63	1.85	0.31	0.29
Difference:	0.25***	2.56***	0.42***	0.27***
	(0.03)	(0.23)	(0.05)	(0.05)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Part B

	5	6	7	8
	Household member(s) participate in disaster management committee	Household member(s) participate in VDC/ward committee	Household member(s) participate in WASH committee	Household member(s) participate in (legally registered) cooperative
Intervention group mean	0.62	0.43	0.39	0.33
Comparison group mean	0.12	0.24	0.07	0.18
Difference:	0.50***	0.19***	0.31***	0.16***
	(0.04)	(0.05)	(0.03)	(0.04)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Although the proportion of households participating in community groups is relatively high across our sample, there are significant differences between the project and non-project households. Of the project households, 88 percent reported participating in at least one community group, compared with 63 percent of the comparison households. For the matched sample, the recalled baseline data suggest that approximately 55 percent of households participated in at least one community group (see Appendix 3). It is important to note that, as we explained in Section 4, our analysis controls directly for baseline differences in group participation, using data recalled back to 2010. Thus, the differences we see in Column 1 must have arisen during the course of the project. This should increase our confidence that the project has had a causal impact on community group participation.

As would be expected, the largest differences between the intervention and comparison groups arise in terms of WEGs and disaster management committees. In particular, the lists of WEG members were used to construct the sample of households in the project areas. As we see in Column 3, 74 percent of the participant households in our matched sample reported having a member involved in a WEG, compared to 31 percent of the comparison households.

Although the difference between the intervention and comparison groups is large, we might, in fact, expect it to be even larger. Firstly, the results from Column 3 indicate that 26 percent of project participant households did *not* report participating in WEGs. In part, our results may underestimate the 'true' level of participation in WEGs. The individuals interviewed may have been unaware of the WEG member within their household. Another possibility is that respondents answered 'No' to avoid follow-up questions especially as the questions about group participation were towards the end of the survey. It is, however, also possible that the WEG membership lists provided by project partners were not completely up to date, and that some people on those lists were not currently participating in WEGs.

Secondly, the fact that 31 percent of the comparison households reported participating in a WEG indicates that women's groups, similar to those set up by the project, operated in the non-project VDCs.[11] Thus, in our analysis, we are not comparing project participant households to a situation in which there were no WEGs or women's groups of any kind. It is important to bear this in mind when interpreting the estimated effects of the project.

It is also worth noting that the higher levels of group participation among the project participant households observed in Table 6.1 are not restricted to the WEGs, which formed the basis of our sampling strategy. Indeed, there are particularly striking differences in terms of participation in disaster management committees and WASH committees, mainly because such a small proportion of the comparison households participated in these types of community groups. In part, this may be because there is substantial overlap between WEG membership and WASH/disaster risk management committee membership. However, the results in Table 6.1 are also likely to reflect the fact that the project sought directly to build up the membership WASH and disaster risk management committees – it is unlikely that the differences we observe are purely a product of our sampling strategy.

In Tables 6.2–6.4, we show a series of results indicating the types of activities that were undertaken in communities by external actors. Respondents were asked whether certain activities had been carried out in their communities since 2010, that is, over the entire duration of the project.

In Table 6.2, we begin by presenting the results for those activities related to stockpiling back-up resources. Among the comparison group, the proportion of households reporting that stockpiling was carried out in their community was very low. In fact, if we run a separate set of statistical tests to determine whether the average

(mean) proportion of household experiencing these activities was significantly different from *zero*, we are unable rule out the possibility that *no one* in the comparison communities was aware of any stockpiling. Given the presence of the ODF and PAF initiatives this is somewhat surprising.

Table 6.2: Stockpiling of back-up resources

	1	2	3
	Stockpiling of warm clothes/blankets	Stockpiling of back-up food supplies	Stockpiling of tarpaulin
Intervention group mean	0.33	0.30	0.29
Comparison group mean	0.01	0.01	0.01
Difference:	0.32***	0.29***	0.28***
	(0.03)	(0.03)	(0.03)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Among the project participants, by contrast, around 30 percent of the households reported that each type of stockpiling activity was undertaken in their community. Therefore, even though stockpiling was far more prevalent in the intervention group than the comparison group, it seems that either these stockpiling activities were by no means ubiquitous in the project VDCs or respondents in the intervention group were unaware that they were happening.

Table 6.3 shows the differences between intervention and comparison households in terms of the training in which they participated during the life of the project. There are positive and significant differences in favour of the project across training on first aid, search and rescue, and WASH. The largest differences arise for search and rescue training, which was virtually absent from the comparison group, but received by around half of the project households in our sample. Interestingly, although the project nearly doubled the proportion of households that had received WASH training, 45 percent of non-project households had been trained on WASH practices. This fits with our sampling strategy, which sought to focus on VDCs where the government's ODF initiative was being implemented (for both intervention and comparison communities). Again, however, this should be borne in mind when interpreting the results that follow – we are not simply comparing the project households to a situation without any WASH support at all.

Table 6.3: Trainings

	1	2	3
	First aid training	Search and rescue training	WASH training
Intervention group mean	0.57	0.49	0.81
Comparison group mean	0.12	0.01	0.45
Difference:	0.44***	0.48***	0.36***
	(0.04)	(0.03)	(0.04)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Finally, we report the differences between project and comparison VDCs in terms of direct construction projects in Table 6.4 (based on households' responses). The proportion of households reporting that each of these construction activities had been carried out in their community was significantly higher in the intervention group than the comparison group. These results persist across construction projects related to flood mitigation (Columns 1 and 2), land improvement (Columns 3 and 4), and WASH (Columns 5 and 6). However, there appears to be some variation across types of construction work in terms of prevalence in the comparison group. In particular, more than 20 percent of comparison households reported that WASH facilities – such as wells, boreholes, and toilets – had been constructed in their community. This, once again, reinforces the notion that government programmes, such as the ODF and PAF, were providing some assistance, especially on WASH, to the non-project households in our sample.

Table 6.4: Construction activities

Part A

	1	2	3
	Small-scale flood mitigation structure (e.g. embankment)	Evacuation route (including culvert)	Tree-planting
Intervention group mean	0.51	0.42	0.44
Comparison group mean	0.08	0.12	0.09
Difference:	0.43***	0.30***	0.36***
	(0.04)	(0.04)	(0.04)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Part B

	4	5	6
	Irrigation system	Wells/boreholes	Toilets/other WASH facilities
Intervention group mean	0.49	0.43	0.68
Comparison group mean	0.20	0.20	0.26
Difference:	0.28***	0.23***	0.42***
	(0.05)	(0.05)	(0.04)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Therefore, taking the results from Section 6.2 together, the intervention group clearly participated in more activities that were implemented by the project than the comparison group. This is as expected. However, project households did *not* universally report having participated in the activities or having received the various types of support shown in Tables 6.1–6.4. In the comparison group, only a very small proportion of respondents reported receiving external support directly related to disaster risk management. However, a substantial minority of comparison households reported that activities to support WASH were carried out in their communities, as would be expected based on our decision to focus on VDCs (for both the intervention and comparison group) that were participating in the government's ODF and PAF initiatives.

6.3 LIVESTOCK

In this section, we examine how the project influenced households' ownership of livestock. Livestock not only provides households with alternative sources of food and income – through, for example, the production of milk and eggs – but also enables households to undertake agricultural activities more effectively – by helping to plough fields, transport agricultural output, and so on. Additionally, livestock can be sold to support incomes directly.

In Table 6.5 we present the differences between the project and non-project households in terms of overall livestock ownership. The results in Column 1 suggest that the project has had a small but statistically significant effect on whether or not households owned livestock. Approximately 82 percent of project households owned livestock at the time of the survey, compared with 72 percent of the comparison households.[12] Thus, the vast majority of sampled households owned some type of livestock. In Column 2 we can also see that there were moderate differences in terms of the number of types of animals owned in project and non-project households.

Table 6.5: Overall livestock ownership

	1	2	3	4	5
	Household owned any livestock	Number of types of animals owned	Women mainly responsible for any types of animal	Number of animal types for which women were mainly responsible	Proportion of animal types for which women were mainly responsible
Intervention group mean	0.82	1.66	0.28	0.42	0.30
Comparison group mean	0.72	1.23	0.21	0.30	0.27
Difference:	0.10**	0.43***	0.07	0.12*	0.03
	(0.04)	(0.10)	(0.04)	(0.07)	(0.05)
Observations (intervention group)	272	272	272	272	223
Observations (total)	799	799	799	799	581

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

In Columns 3, 4, and 5 we explore whether women in project households had more responsibility for looking after livestock than comparison households. For each type of animal owned, households were asked which household members were mainly responsible for looking after that animal – these results relate to those animals for which female members were 'mainly responsible'. However, in general, the project did not appear to have statistically significant effects on women's responsibility for livestock. Across Columns 3, 4, and 5 we report whether or not women were mainly responsible for any animal types, the number of types for which they were responsible, and the proportion of animal types owned for which they were responsible (respectively). The only statistically significant differences arise in Column 4 for the number of types of animals for which women were mainly responsible. However, this may simply reflect the fact that project households had a more diversified livestock portfolio anyway, rather than telling us about women's responsibility for household assets.

In Table 6.6 we report the *number* of each type of animal owned by project and non-project households. Before considering the differences, it should be noted that even though three-quarters of households across the entire sample own some livestock, the number of livestock they own appears to be quite limited – this explains why so many of the numbers in Table 6.6 are less than one. Indeed, even if we restrict the sample to those households that actually own goats (for example) the average number of goats owned is just three. Thus, it seems that for most households in the sample, livestock

provides a supplementary livelihood activity, rather than providing their main source of income.

Table 6.6: Numbers of each animal owned

	1	2	3	4	5	6
	Cows	Buffalo	Oxen	Goats	Pigs	Poultry
Intervention group mean	0.72	0.58	0.43	1.76	0.08	0.76
Comparison group mean	0.35	0.37	0.26	1.42	0.01	1.28
Difference:	0.37***	0.21**	0.17***	0.34	0.07	-0.52
	(0.14)	(0.10)	(0.06)	(0.21)	(0.05)	(0.99)
Observations (intervention group)	272	272	272	272	272	272
Observations (total)	799	799	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Nonetheless, the project appears to have had a positive and statistically significant impact on the ownership of large animals – cows, buffalo, and oxen. It turns out that this reflects both an increase in the likelihood of owning *any* large animals and an increase in the *number* of large animals (among those households that owned large animals anyway). By contrast, there were no significant differences between the project and comparison households for the small animals, even for goats (the most common type of livestock in this context).

Therefore, the results for livestock ownership are somewhat mixed. The project led to a modest increase in the likelihood that households owned livestock as well as the number of types of animals owned. However, there was no clear evidence that the project raised women's responsibility for animals within the household. Also, increases in livestock ownership appear to have been concentrated in large animals.

6.4 CROPS

In this section, we consider the project's impact on farming practices. We first outline the project's overall impact on crop growing, highlighting in particular how women's responsibility for crop cultivation and crop sales changed. We then focus directly on whether project households practised more resilient crop growing than their counterparts in the comparison group, especially in terms of off-season crop growing, cultivation of special varieties, and the investments in field fertility.

We start, in Table 6.7, by reporting households' overall engagement in agricultural activities. Approximately three-quarters of the sample engaged in some kind of crop growing in the year prior to the survey, but there were no statistically significant differences between the project and non-project households in the sample, as shown in Column 1. However, the project appears to have had a positive effect on the amount of land households were cultivating. Project households cultivate approximately 2.5 more katthas than their non-project comparators.[13] It is important to note that these results relate specifically to the cultivation of land and not the ownership of land. Indeed, in the questionnaire, the relevant question specifically encouraged respondents to report on land that was rented from other parties or simply occupied, as well as formally owned by the household.

Table 6.7: Overall engagement in agricultural activities

	1	2	3	4
	Household grew any crops in the past year	Number of katthas of land household cultivated	Number of types of crops the household grew in the past year	Number of types of non-cereal crops the household grew in the past year
Intervention group mean	0.75	10.16	2.49	1.10
Comparison group mean	0.69	7.64	1.77	0.46
Difference:	0.06	2.52***	0.72***	0.64***
	(0.04)	(0.88)	(0.16)	(0.12)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

The project also appears to have had a moderate increase on the number of types of crops households cultivated. Project households grew around 2.5 different types of crops, on average, compared to 1.8 crop types in the comparison group. These results persist not just for the crop portfolio in general (Column 3), but even if we focus only on non-cereal crops (Column 4). It should also be noted that this positive result is not changed if we restrict the sample to those crop-growing households.

In Table 6.8 we consider how the project affected women's responsibility for cultivating crops. We first show whether or not women in the household were responsible for farming *any* crops, and then consider the proportion of crop types for which they were responsible. The first point to note is that, both with and without the project, women's responsibility for crop cultivation was fairly limited. Taking the sample of crop-growing households as a whole, women were mainly responsible for just 10 percent of the different types of crops grown (on average). Nonetheless, Column 2 reveals that there was a small difference – around 6 percentage points – between the project and comparison groups in terms of the proportion of types of crops for which women were mainly responsible.[14]

Table 6.8: Women's responsibility for crop cultivation

	1	2
	Women in the household were mainly responsible for farming at least one type of crop	Proportion of types of crops for which women were mainly responsible
Intervention group mean	0.12	0.13
Comparison group mean	0.08	0.08
Difference:	0.04	0.06**
	(0.03)	(0.03)
Observations (intervention group)	272	204
Observations (total)	799	494

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

In Table 6.9, we investigate crop sales. For each type of crop grown in the previous year, respondents were asked what proportion, if any, they sold to market. We report whether or not households sold *any* crops that they grew, and the proportion of crop

types they sold. The project appeared to have positive and statistically significantly effects on these measures. In the intervention group, 25 percent of households sold some crops, approximately double the figure (13 percent) in the comparison group. Among the sub-sample of crop-growing households, there was nearly a two-fold difference in the proportion of crop types sold, as shown in Column 2.

Table 6.9: Crop sales

Part A

	1	2
	Household sold any crops grown	Proportion of types of crops household sold
Intervention group mean	0.25	0.19
Comparison group mean	0.13	0.11
Difference:	0.12***	0.08**
	(0.04)	(0.03)
Observations (intervention group)	272	204
Observations (total)	799	494

Part B

	3	4
	Household sold the entire harvest of any of the crops grown (cash crop)	Proportion of types of crops that are cash crops
Intervention group mean	0.04	0.03
Comparison group mean	0.02	0.02
Difference:	0.02	0.01
	(0.01)	(0.01)
Observations (intervention group)	272	204
Observations (total)	799	494

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

In Columns 3 and 4 we consider whether the project affected the cultivation of ‘cash crops’, that is, those crops for which the harvest was entirely sold. However, the proportion of households growing any of these types of crops was very low, at just 2.7 percent for the sample as a whole. Moreover, there were no statistically significant differences between the intervention and comparison groups.

In the remainder of this section, we examine whether project households cultivated crops in more resilient ways than the non-project households in our sample. We begin, in Table 6.10, by considering whether households grew crops in the off-season or cultivated improved varieties of cereal crops. In Column 1, we report the proportion of households that engaged in off-season farming of any crops in their portfolio. Nearly a quarter of project households cultivated crops in the off-season, compared with a little under 10 percent in the comparison group. This corresponds to a difference of nearly double.

Table 6.10: Off-season crop growing and drought-resistant varieties

	1	2	3	4
	Household grew any crops in the off-season	Household cultivated drought-resistant rice	Household cultivated drought-resistant wheat	Household cultivated drought-resistant maize
Intervention group mean	0.23	0.36	0.38	0.04
Comparison group mean	0.09	0.19	0.24	0.03
Difference:	0.13***	0.17***	0.14***	0.01
	(0.03)	(0.04)	(0.04)	(0.02)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

In the remaining columns of Table 6.10, we report the proportion of households that had grown improved varieties of cereal crops. Respondents were asked whether their household had cultivated 'drought-resistant' types of rice, wheat or, maize in the previous year. We show these results for the sample as a whole, meaning that the proportion of households *not* growing improved varieties of rice, say, includes both those households growing unimproved rice varieties and no rice at all. There appear to be positive and significant differences between the intervention and comparison groups in terms of drought-resistant rice and drought-resistant wheat growth. In project VDCs, 36 percent of households grew drought-resistant rice varieties, compared with just 19 percent of the non-project households in our sample. This corresponds to a difference of nearly double the figure for the comparison group. Similarly, the proportion of households cultivating drought-resistant wheat was a little under 60 percent higher for the project participants. These results do not change substantially if we restrict the sample to those households who grew *any* type of rice or wheat. However, there are no statistically significant differences between project and non-project households in term of growing drought-resistant maize. In part, this may be because the proportion of households growing any type of maize in this context is very low – maize-growing households comprise just 12 percent of the sample.

The final set of results we report in this section, shown in Table 6.11, examine the project's effects on field fertility. Here, we restrict the sample to those households that were actually cultivating any types of crops in the year prior to the survey.

Table 6.11: Field fertility

	1	2	3	4	5
	Field fertility increased, in the last two years	Household irrigates some proportion of its agricultural land	Household irrigates more than 50 percent of its agricultural land	Household uses fertiliser (any type) for its agricultural activities	Household uses vermi-compost for its agricultural activities
Intervention group mean	0.22	0.71	0.47	0.97	0.13
Comparison group mean	0.03	0.69	0.37	0.94	0.12
Difference:	0.19***	0.03	0.10*	0.02	0.02
	(0.03)	(0.06)	(0.06)	(0.02)	(0.03)
Observations (intervention group)	204	204	204	204	204
Observations (total)	494	494	494	493	493

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

In Column 1, we show the proportion of households reporting that the fertility of their fields had increased over the *two* years prior to the survey. Around 22 percent of the project households said they experienced such an improvement, compared to just 3 percent of the comparison group. Although this difference is statistically significant, it is worth noting that the vast majority of intervention group households (78 percent) did not report a positive change in the fertility of their fields. In any case, we should be somewhat cautious about placing too much weight on the results in Column 1 because this questionnaire asked respondents to provide a *subjective* rating of field fertility. Thus, our results may be biased if respondents systematically under- or over-reported changes in the fertility of the fields. Moreover, these results tell us only about changes over time, and do not provide direct information about the current status of respondents' fields.

Therefore, in Columns 2 to 5 of Table 6.11 we present the results from some more objective and static indicators of field fertility. In Columns 2 and 3 we consider first if the project influenced whether or not households irrigated *any* of their land, then second whether or not the household irrigated the *majority* of their land. Although there are no statistically significant effects for the first metric, 47 percent of project households irrigated the majority of their land compared to just 37 percent of the non-project households (and this difference is statistically significant at the 10 percent level).

In the final two columns, of Table 6.11 we observe that almost all farming households were using fertiliser, and that there are no statistically significant differences between the intervention and comparison groups in this respect. Indeed, even if we focus only on vermi-compost, a specific project activity, it does not appear that the project had any significant effect on take-up.

Thus, the overall impact on crop growing is positive, but with some caveats. The project appears to have increased the amount of land households farmed, and helped households cultivate a wide range of crop types. Women in project households had slightly more responsibility for the crop portfolio, on average, than non-project households, and while project households were more likely to sell some crops, they were no more likely to cultivate cash crops. In terms of resilient crop cultivation practices, the strongest positive results arose in terms of off-season farming, and cultivation of drought-resistant cereal varieties.

6.5 NON-FARM ACTIVITIES

In this final section on livelihood strategies, we consider the Joint Programme's impact on non-farm income-generating activities. As discussed in Section 2, this was a key avenue through which the project aimed to increase resilience.

In Table 6.12, we begin (in Column 1) by showing the proportion of households engaging in non-farm activities in the intervention and comparison groups. If anything, project households were *less* likely to engage in non-farm livelihood strategies than the non-project households in our sample. Approximately 59 percent of households in project VDCs engaged in non-farm income-generating activities, compared with 67 percent of the comparison VDCs. This difference was significant at the 5 percent level.[15] There were no statistically significant differences between the intervention and comparison groups in terms of the number of non-farm income-generating activities in which households engaged, and the proportion of households where female members engaged in non-farm livelihoods, as shown in Columns 2 and 3. It is, however, interesting to note how low the proportion of households containing women engaged in non-farm income-generating activities is across the sample as whole. Taking the intervention and comparison groups together, just 17 percent of households contained women engaged in non-farm activities.

Table 6.12: Non-farm livelihood activities

Part A

	1	2	3
	Household engaged in any non-farm income-generating activities in the past 12 months	Number of non-farm income-generating activities that the household engaged in	Female household member engaged in any non-farm income-generating activities
Intervention group mean	0.59	0.75	0.17
Comparison group mean	0.67	0.82	0.18
Difference:	-0.08**	-0.08	-0.01
	(0.04)	(0.07)	(0.04)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Part B

	4	5	6
	Casual labour (e.g. construction, working short-term in a factory)	Household business (e.g. retail shop, hotel, tea shop)	Regular, paid formal job (e.g. teacher, nurse, public sector, long-term factory
Intervention group mean	0.49	0.13	0.13
Comparison group mean	0.58	0.13	0.12
Difference:	-0.09**	-0.00	0.02
	(0.05)	(0.04)	(0.04)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

In Columns 4, 5, and 6 of Table 6.12 we disaggregate the results by considering whether households' likelihood of participating in each different off-farm occupation was affected by the project. We can see that the results in Column 1 are being driven by casual labour, such as construction or short-term work in a factory (reported in Column 4). Approximately 49 percent of project households undertook this type of casual work, compared with 58 percent of the comparison group. Without knowing more about the specific jobs included in casual work, it is impossible to say whether this is a positive or a negative change. Although having more diversified income sources is positive, all other things being equal, it may be that casual labour has low returns relative to farming, especially if the project has successfully improved agricultural productivity.

We do, however, see that there are no significant differences between the intervention and comparison group in terms of engaging in non-farm household businesses, and regular paid formal work. These results persist, even if we restrict the analysis to women in the household. The apparent lack of impact on household businesses is especially puzzling, as this was a specific aim of the WEGs. In part, this may be because the project activities that were aimed at supporting livelihoods were only implemented in the project's later stages.

Overall, therefore, our data do not provide evidence of positive change for non-farm income generating activities.

6.6 WASH

At various points during the questionnaire, respondents were asked about their WASH practices. In this section, we present some of the key results from these questions.

Firstly, the livestock-owning households in our sample were asked where they stored their animals in relation to the main kitchen in their home. Maintaining distance between the cooking area and the household's animals is thought to demonstrate good hygiene practice. The proportion of households storing their livestock in particular ways is shown in Table 6.13 – it should be noted that these results are restricted to the sub-sample of households that owned any animals at the time of the survey.[16]

Table 6.13: Livestock storage practices

	1	2	3
	Same room	In different structure	Livestock kept outside with no shed
Intervention group mean	0.19	0.34	0.04
Comparison group mean	0.34	0.20	0.12
Difference:	-0.14***	0.14***	-0.07**
	(0.05)	(0.05)	(0.03)
Observations (intervention group)	226	226	226
Observations (total)	600	600	600

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

It emerges that a smaller proportion of project households stored their livestock in the main kitchen itself, compared with the non-project households in our sample (Column 1). Approximately 19 percent of the intervention group kept their livestock in the same room that they prepared food, compared with 34 percent of the comparison group. Although this measure is insensitive to the size and shape of this room, this suggests that WASH practices around livestock were improved by the project. This is reinforced by the fact that significantly more project households kept their livestock in a totally different structure than the comparison households. In Column 2, we see there is a difference of around 14 percentage points. In addition, as Column 3 shows a significantly smaller proportion of the intervention group (4 percent) kept their animals outside with no shed compared with the comparison group (12 percent). Cordoning livestock off to prevent them from roaming around different areas of the household may not only improve WASH practices, but also may help protect livestock from floods and cold waves.

We next move to consider whether the project had any impact on the main source of drinking water that the household used at the time of the survey. Here the results, which are shown in Table 6.14, are more ambiguous. Firstly, there are no clear differences between the intervention and comparison groups in terms of the proportion of households that typically take their drinking water from a tap in the house and from a well/borehole. However, there appear to be more project households using water sources such as ponds, rivers, and streams. Approximately 9 percent of the intervention group used these types of sources for drinking water, compared with just 4 percent of the comparison group. It should be noted, however, that there were similar differences between the project and non-project households in 2010, according to the recalled baseline data, even after matching. Therefore, the results in Table 6.14 may reflect existing differences between the intervention and comparison groups, which we were not able to accommodate in our matching procedure, rather than causal effects of the project.

Table 6.14: Main water source of drinking water for the household

	1	2	3
	A tap in the house	Well/borehole	Other (e.g. pond, river, stream)
Intervention group mean	0.07	0.84	0.09
Comparison group mean	0.09	0.87	0.04
Difference:	-0.02	-0.04	0.06**
	(0.03)	(0.03)	(0.03)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

6.7 RISK MANAGEMENT STRATEGIES

In this section, we briefly report the adoption of coping and risk-management strategies in the sample. Households were asked whether they had taken a list of actions to protect themselves, their crops, or their assets from disaster during the two years prior to the survey. The proportion of households reporting that they had undertaken each of these actions is shown in Table 6.15.

Table 6.15: Coping strategies

	1	2	3	4	5
	Changing cropping patterns	Improving area where livestock/ assets are kept for more protection	Storing grain/ crops in a special container (e.g. grain bin)	Improving the structure of the house	Raising household (e.g. on stilts, special plinth, or building second storey)
Intervention group mean	0.32	0.38	0.54	0.38	0.35
Comparison group mean	0.03	0.04	0.34	0.09	0.09
Difference:	0.28***	0.33***	0.20***	0.29***	0.26***
	(0.03)	(0.03)	(0.05)	(0.04)	(0.04)
Observations (intervention group)	272	272	272	272	272
Observations (total)	799	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

The overall picture emerging from the results is that these risk-management strategies were undertaken more commonly among the project households than the non-project households in our sample. Some of these findings relate to previous results in the report. For example, we have already seen evidence that the project may have supported households to change crop patterns and improve the area where they keep their livestock. These results are supported by Columns 1 and 2.

The results in Table 6.15 also highlight that project households have undertaken a number of investments to better protect their homes and their agricultural activities. For example, in Columns 4 and 5, we see that 38 percent of project households had made improvements to the structure of their house, and 35 percent had even attempted to raise their houses to better protect them from floodwater. The analogous proportions in the comparison group were under 10 percent.

We should, however, be somewhat cautious about interpreting the results in Table 6.15 in a completely positive light. In particular, it may be that these measures were necessary due to project households' continued exposure to risk. This is especially concerning if these sorts of investments come at the expense of investing in income-generating activities. Given the equivocal results on livelihood strategies in Section 6.5, this issue should certainly be borne in mind.

6.8 WEALTH

In this section, we explore the project's impact on households' wealth. Wealth may be interpreted in two ways from the perspective of resilience. Firstly, wealth may be seen as a *driver* of resilience, insofar as households can sell off assets in times of crisis but also more easily finance the costly investments needed to adapt livelihood strategies and innovate. However, wealth may also been regarded as exactly the type of wellbeing indicator – a 'final' outcome – which would be improved in spite of shocks, stresses, and uncertainty in more resilient households. Typically, these types of final wellbeing outcomes take more time to change than more immediate drivers or characteristics of resilience.

During the course of the questionnaire, respondents were asked to provide information about their household's ownership of various assets (including livestock, productive equipment, and household goods), as well as about the conditions of the family's house, both in 2010 and at the time of the survey. This information on asset ownership and housing conditions was used to generate an index of overall household wealth.

The wealth index was generated under the assumption that if each of the assets and housing characteristics constituted suitable indicators of household wealth, they should be correlated with each other. That is, a household that scores favourably on one particular wealth indicator should be more likely to do so for other wealth indicators. A small number of items that had low or negative correlations with the others were therefore not considered to be good wealth indicators and so were excluded from the index.[17]

A data reduction technique called principal component analysis (PCA) was used to produce two indices of overall wealth, one based on the recalled data from 2010, and one based on the household's situation at the time of the survey. In particular, our wealth index is taken directly from the first principal component.[18] PCA enables us to assign weights to the different assets, to capture as much information as possible from the data. Broadly, PCA assigns more weight to those assets that are *less* correlated with all the other assets, as these carry more information. By contrast, items with *more* intra-correlation are given less weight.

In order to ensure the same weights were applied to assets for both the recalled wealth index and the wealth index for the time of the survey, data from these two time periods were pooled before undertaking the PCA procedure. This means changes in wealth can be more easily compared over time. It should also be noted that the wealth index for 2010 is the measure that has been used throughout this analysis to control for baseline differences in wealth status between project and non-project households.

For the analysis in this section, we start by 'normalising' the wealth index.[19] This means that the impacts of the project that we report can be directly understood as the *number* of standard deviations by which the project improved wealth. This means the results from this Effectiveness Review can be more easily compared to other similar evaluations.

In Table 6.16, we estimate the project's impact on wealth in two ways. In Column 1 we report wealth for the project and non-project households at the time of the survey,

using the regular matching procedure that has been used throughout the other tables in this report. In Column 2, however, we take a slightly different approach. We calculate the differences between wealth at the time of the survey and in 2010, and compare these differences between project and non-project households in the matched sample. For the results in Column 2, it is necessary to omit recalled wealth from the matching process.[20]

Table 6.16: Wealth

	1	2
	Normalised wealth index	Difference in normalised wealth index
Intervention group mean	0.61	0.88
Comparison group mean	0.53	0.76
Difference:	0.07	0.12
	(0.11)	(0.08)
Observations (intervention group)	272	272
Observations (total)	799	791

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Although both the current wealth index and the difference in the wealth index were higher in the project households than the comparison households, these differences are not statistically significant.[21] Therefore, our data do not provide evidence that the project had a positive impact on household wealth. It should, however, be remembered that the survey work for this evaluation took place before the project had finished. Changes in wealth status may require a longer time horizon in this context. As such, it may be useful to follow up this evaluation with future efforts to measure wealth status after a number of years.[22]

6.9 INDICATORS OF RESILIENCE

In Section 5 we outlined our approach for measuring resilience in this Effectiveness Review. We described how to construct an overall index for resilience, with four potential strategies for weighting indicators against one another. The project's impact on resilience, measured in this way, is shown in Table 6.17.

Table 6.17: Indices of resilience

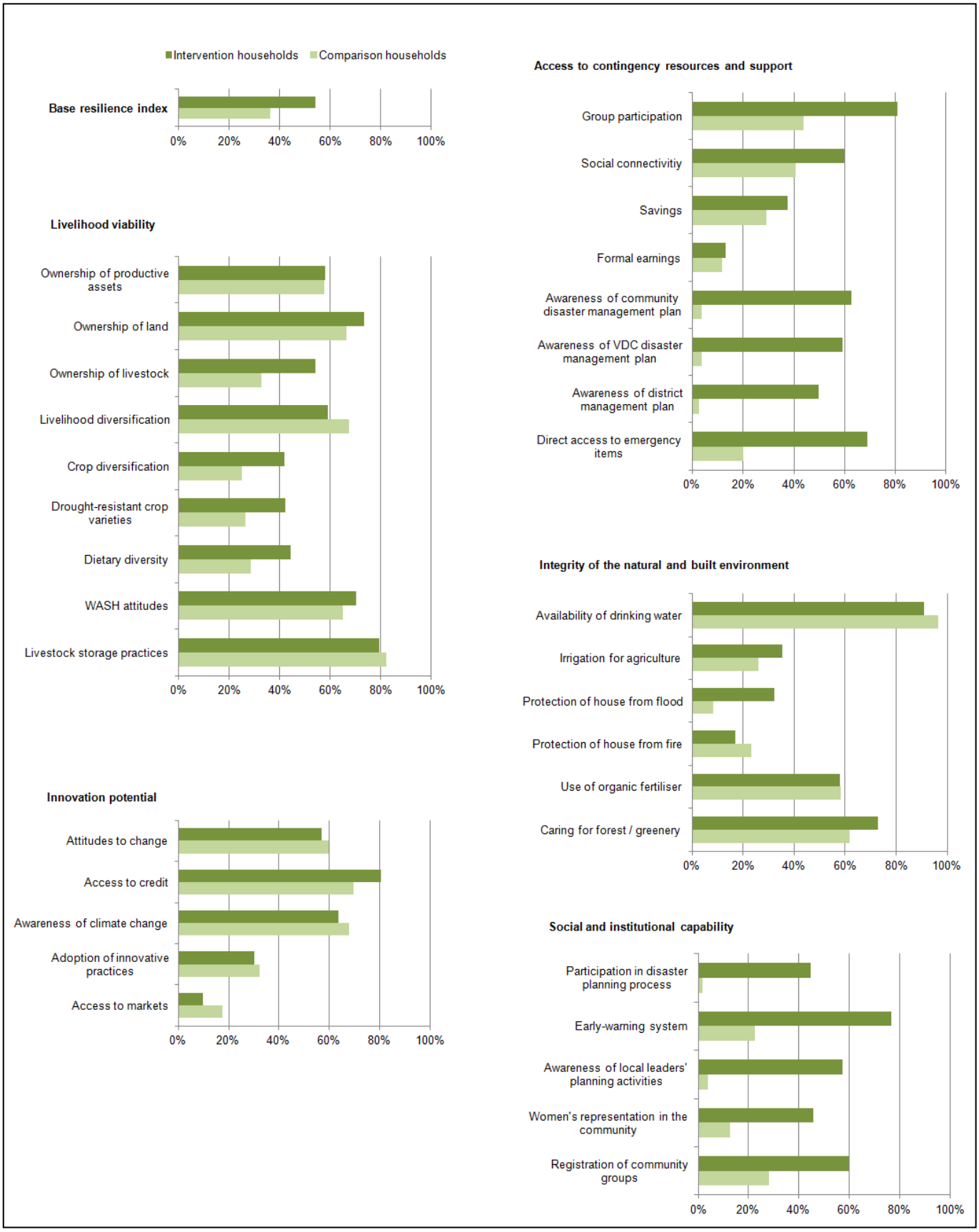
	1	2	3	4
	Base resilience index	Equal dimensions resilience index	Personal dimensions resilience index	Sample dimensions resilience index
Intervention group mean	0.53	0.53	0.53	0.53
Comparison group mean	0.34	0.34	0.33	0.32
Difference:	0.18***	0.18***	0.20***	0.21***
	(0.01)	(0.01)	(0.01)	(0.01)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

There is clear evidence that the project increased households' resilience. Using all four weighting strategies, the resilience index was between 18 and 21 percentage points higher in the intervention group compared with the comparison group. Focusing just on the base resilience index in Column 1, these results imply that project households scored positively, on average, in 53 percent of the resilience indicators, compared with just 34 percent for the comparison households. This means the base resilience index was nearly 50 percent higher among the project households, using the non-project households in the matched sample as a reference point. All of these differences were statistically significant, even at the 1 percent level.

To understand better what is driving these increases in the resilience index, it is important to compare project and comparison households in terms of all the constituent indicators. To gain an initial overview, Figure 6.18 presents the average proportions of intervention and comparison group households scoring positively on each of the resilience indicators. We then describe the results under each dimension of resilience in the sub-sections that follow.

Figure 6.18: Results for characteristics of resilience



6.9.1 Dimension 1: Livelihood viability

Nine indicators of resilience falling under the 'livelihood viability' dimension were examined in this Effectiveness Review. The project's impact on these characteristics is shown in Table 6.19. Since the construction of each indicator is explained in Appendix 1, we focus here on the stand-out results.

Table 6.19: Indicators of livelihood viability

Part A

	1	2	3	4	5
	Ownership of productive assets	Ownership of land	Ownership of livestock	Livelihood diversification	Crop diversification
Intervention group mean	0.58	0.74	0.54	0.59	0.42
Comparison group mean	0.58	0.66	0.33	0.67	0.25
Difference:	0.00	0.07*	0.21***	-0.08**	0.17***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Observations (intervention group)	272	272	272	272	272
Observations (total)	799	799	799	799	799

Part B

	6	7	8	9
	Drought-resistant crop varieties	Dietary diversity	WASH attitudes	Livestock storage practices
Intervention group mean	0.42	0.44	0.70	0.28
Comparison group mean	0.27	0.28	0.65	0.16
Difference:	0.16***	0.16***	0.05	0.13***
	(0.04)	(0.05)	(0.04)	(0.04)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Many of the results in Table 6.19 echo findings that have been previously discussed in this report. For example, we see that the project households were more resilient in terms of livestock ownership, crop diversification, and cultivation of drought-resistant crops. In contrast, the fact that livelihood diversification was lower in the intervention group than the comparison group is also shown in Column 4 of Table 6.19.

There are two further positive results in Table 6.19, which have not been highlighted already in the report. Firstly, as we see in Column 7, project households outperform non-project households in terms of dietary diversity. On this measure, households scored positively if, in the last 7 days, they had consumed a carbohydrate source every day, a protein source on at least three days, and some fruit or vegetables on at least three days. Approximately 44 percent of project households met this condition, compared with 28 percent of the comparison group. This 16 percentage point gap corresponds to nearly a 60 percent difference, using the comparison group as a reference point. In part, this emphasises our results for crop diversification and livestock ownership, insofar as households are able to consume their own more varied agricultural produce. Dietary diversity may also reflect household members' nutritional intake, proxying for their resilience to health-related shocks. One final possibility is that households with more diverse diets are better integrated with markets, as they are able

to source a wide range of foods. There are, therefore, a number of lenses through which to interpret the positive result shown in Column 7.

There is also a modest difference between the intervention and comparison groups in terms of ownership of land, presented in Column 2. Households scored positively on this measure if they legally owned *any* land at all, with the *lal purja* document to back up their claim. As Column 2 shows, 74 percent of project households legally owned land, compared with 66 percent of the comparison group, an increase of just over 10 percent. Although this difference is only statistically significant at the 10 percent level, it is important to note (as mentioned in Section 4) that our analysis controls directly for baseline differences in land ownership, using wealth data recalled from 2010. This should increase our confidence that the differences we observe in Column 2, are truly down to the effects of the project and do not result from unobservable differences between the intervention and comparison group. Indeed, we also showed in Appendix 3 that there were no statistically significant differences in terms of land ownership in 2010 between the project and non-project households in the matched sample, according to the recalled baseline data.

Table 6.19 also contains some equivocal results, where it was not possible to identify clear differences between the intervention and comparison groups with our data. For the ownership of productive assets (Column 1) and WASH attitudes (Column 8), we cannot say with any confidence whether the project had either a positive or a negative impact.

6.9.2 Dimension 2: Innovation potential

We identified five characteristics that were thought to capture 'innovation potential' in this Effectiveness Review. We report these results in Table 6.20, once again concentrating on the main results in the following text and reserving the full explanation of each indicator for Appendix 1.

Table 6.20: Indicators of innovation potential

	1	2	3	4	5
	Attitudes to change	Access to credit	Awareness of climate change	Adoption of innovative practices	Access to markets
Intervention group mean	0.57	0.81	0.64	0.30	0.10
Comparison group mean	0.60	0.70	0.68	0.32	0.17
Difference:	-0.03	0.11***	-0.04	-0.02	-0.08*
	(0.05)	(0.03)	(0.04)	(0.05)	(0.04)
Observations (intervention group)	272	272	272	272	272
Observations (total)	799	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Project households had better access to credit than their non-project comparators. Households were scored positively in terms of access to credit if they would be able to borrow 10,000 rupees to invest in a business opportunity from someone other than friends, family or, a local moneylender. Approximately 81 percent of project households reported that they had this level of access to credit, compared to just 70 percent of non-project households. This matches the intended outcome described in the project logic in Section 2. However, it is unclear why these relaxed credit constraints appear to have mainly facilitated more investment in and experimentation with agriculture, but not off-farm livelihood strategies.

By contrast, the remaining results in Table 6.20 do not show positive evidence in favour of the project. There were no statistically significant differences in terms of attitudes to

change (Column 1), awareness of climate change (Column 3) and adoption of innovative practices (Column 4). It should be noted that attitudes to change and awareness of climate change are measured through opinions questions, which asked respondents to choose between two opposing statements. For example, for attitudes to change, respondents were asked whether they agreed more with (1) 'We should not be afraid to try new and different livelihood activities – sometimes they are better than the traditional livelihood activities' or (2) 'It is best to continue doing what we already know and do well, rather than experimenting with new approaches'. Those households agreeing with option (1) scored positively in terms of attitudes to change. However, given that self-perception measures of this type may be susceptible to certain biases, we should be cautious about emphasising these results too much.

In the final column of Table 6.20, we see that the project appeared to have a negative impact on households' access to markets. Specifically, households scored positively on this indicator if they experienced no difficulties reaching the nearest market during the monsoon season. The results in Column 5 mean that 90 percent of project households experienced at least some of these types of difficulties compared with 83 percent of the non-project households, although this difference is only statistically significant at the 10 percent level). This may help explain why the project's positive effects on credit constraints have not translated into the establishment of household businesses and other non-farm income-generating activities.

6.9.3 Dimension 3: Access to contingency resources and support

We used eight indicators to measure 'access to contingency resources and support' in this Effectiveness Review. These results are presented in Table 6.21.

Table 6.21: Indicators of access to contingency resources and support

Part A

	1	2	3	4
	Group participation	Social connectivity	Savings	Formal earnings
Intervention group mean	0.81	0.60	0.38	0.13
Comparison group mean	0.44	0.41	0.29	0.12
Difference:	0.37***	0.19***	0.08*	0.02
	(0.04)	(0.05)	(0.05)	(0.04)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Part B

	5	6	7	8
	Awareness of community disaster management plan	Awareness of VDC disaster management plan	Awareness of district management plan	Direct access to emergency items
Intervention group mean	0.63	0.59	0.50	0.69
Comparison group mean	0.04	0.04	0.03	0.20
Difference:	0.59***	0.55***	0.47***	0.49***
	(0.03)	(0.03)	(0.04)	(0.05)
Observations (intervention group)	272	272	272	272
Observations (total)	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Many of these indicators are directly related to the project logic, so some of the positive results we see in Table 6.21 are unsurprising. For example, we know the project worked directly to raise group participation and to provide direct access to emergency items – such as warm clothes, first aid kits, and life jackets – as Column 8 shows. The results in Columns 5, 6, and 7 indicate a somewhat nuanced picture for the project's impact on disaster management plans. Clearly the efforts to orientate project communities with community-, VDC- and district-level disaster management plans worked well. Awareness of these plans was substantially higher among project households than non-project households. However, the project's work on *district*-level plans does not seem to be reaching non-project VDCs. Among the comparison group, just 3 percent of households reported being at least partly aware of the district-level plans. Therefore, although project communities substantially outperformed their non-project comparators on these metrics, any support being provided at the district-level does not appear to affect the comparison group.

The project also had positive and statistically significant effects on two further indicators of resilience, which were less directly related to the project activities. Firstly, households appeared to have improved social connectivity – this was an intermediate outcome, which was relatively high up in the project's logic model. Households scored positively on this measure if they could rely on friends inside and/or outside the community if they needed to borrow 10,000 rupees to invest in a business opportunity, *and* the respondent reported that they would help their neighbours with food, money, and other commodities during hard times – and believed their neighbours would do the same for them. This may partially reflect the fact that WEGs specifically sought to financially integrate community members.

There is also some tentative evidence that project households had higher savings than non-project households. In Column 3, we see that 38 percent of households in the intervention group had enough money saved to last at least 30 days in an emergency, compared with 29 percent of the comparison group. This difference is statistically significant at the 10 percent level. Again, this supports the idea that the WEGs enabled households to borrow and save more. It should, however, be noted that these results are somewhat sensitive to the cut-off we use, and do not survive if we use a continuous specification (that is, simply the number of days that money savings would cover) as an outcome variable.[23] As such, we cannot say with confidence that the project had a positive impact on savings.

Finally, we note that, as shown in Section 6.5, there are no statistically significant differences between the project and non-project household in terms of regular, formal paid work.

6.9.4 Dimension 4: Integrity of the natural and built environment

Since this Effectiveness Review used household-level data, finding meaningful characteristics of the natural environment was difficult. Nonetheless, we present the evidence from six indicators thought to capture the 'integrity of the natural and built environment' in Table 6.22.

Table 6.22: Indicators for the integrity of the natural and built environment

Part A

	1	2	3
	Availability of drinking water	Irrigation for agriculture	Protection of house from flood
Intervention group mean	0.91	0.35	0.32
Comparison group mean	0.96	0.26	0.08
Difference:	-0.06**	0.09**	0.24***
	(0.03)	(0.05)	(0.03)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Part B

	4	5	6
	Protection of house from fire	Use of organic fertiliser	Caring for forest/ greenery
Intervention group mean	0.17	0.58	0.73
Comparison group mean	0.23	0.58	0.62
Difference:	-0.06	-0.00	0.11**
	(0.04)	(0.05)	(0.05)
Observations (intervention group)	272	272	272
Observations (total)	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

The most striking positive results arise in Column 3, which show that households in project VDCs were better protected from flooding, on average, than the non-project households in our sample. Households were rated as protected if they were artificially raised up, for example on a plinth or special stilts. Approximately 32 percent of project households met this condition, compared with just 8 percent of non-project households. The project did not directly provide support in raising houses, so this finding suggests that the project instead *enabled* participant households to undertake resilience-building initiatives of their own. In order to check that the results in Column 3 do not arise from baseline differences between the intervention and comparison groups, we add recalled data on whether or not households were artificially raised up into the matching function. It emerges that the project's positive effects on household flood protection are not sensitive to adding this information to the set of matching variables. This should reassure us that the results in Column 3 represent positive effects of the project, rather than simply showing that project VDCs were more vulnerable to flooding.

However, as Column 4 shows, there are no such differences between the intervention and comparison group in terms of protecting the household from fire, that is, using non-flammable materials for the walls and roof of the house.

A number of the results in Table 6.22 further emphasise findings that have been presented in previous sections of the report. For example, in terms of sustainable agricultural practices, we see that project households appear to irrigate their land more than comparison households, but that there are no significant differences in terms of organic fertiliser use. By contrast, we also saw that a greater proportion of project households relied upon risky water sources – such as ponds, streams, and rivers – compared with the non-project households in our sample. However, as we noted in Section 6.6, it appears that these differences were present before the project started.

Finally, project households appeared to be more concerned about preserving forests and other natural resources than non-project households, as the results in Column 6 demonstrate. Households scored positively on this measure if respondents reported that it was sometimes necessary to reduce the use of natural resources to ensure they would be available in the future and/or the household participated in a community forest group. Approximately 73 percent of households met this condition, compared with 62 of the non-project households – a difference of 11 percentage points, and a relative increase of 18 percent. It should be noted that, if we disaggregate this indicator into its two constituent parts – attitudes to natural resource use and participation in community forest groups – there are positive and significant differences for *both* of these components. This suggests that in terms of using resources sustainably the project was able to change both people's attitudes and their actions.

6.9.5 Dimension 5: Social and institutional capability

The project worked directly on building 'social and institutional capability', so finding indicators for that dimension was a particularly important step in this evaluation. However, since our data were collected at the household-level, we may be unable to capture fully the elements of social and institutional capability that relate to systemic transformation (such as intra-household bargaining, community-level leadership, district governance, and so on). Nonetheless, we were able to establish five indicators from our data, the results for which are reported in Table 6.23.

Table 6.23: Indicators of social and institutional capability

	1	2	3	4	5
	Participation in disaster planning process	Early-warning system	Awareness of local leaders' planning activities	Registration of community groups	Women's representation in the community
Intervention group mean	0.45	0.77	0.57	0.60	0.46
Comparison group mean	0.02	0.23	0.04	0.28	0.13
Difference:	0.43***	0.54***	0.53***	0.32***	0.33***
	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)
Observations (intervention group)	272	272	272	272	272
Observations (total)	799	799	799	799	799

Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1,000 repetitions.

There are large and statistically significant differences between the project and non-project households in our sample, across all the indicators examined in Table 6.23. Although we have tried to find intermediate outcomes, the results in the first four columns are very much related to the project activities. The project worked directly to improve participation in the disaster planning process, build confidence in early-warning systems, and help community groups register formally with the authorities. Although these indicators are therefore quite low-level in the project logic, we include them because they were mentioned not only during discussions with project staff, but also during focus groups with local communities. Again, the striking feature of these results is not so much the difference between the project and non-project households, but rather the especially low proportion of households in the comparison group reporting that they (1) had participated in the disaster planning process (Column 1), and (2) were aware of local leaders' planning activities (Column 3).

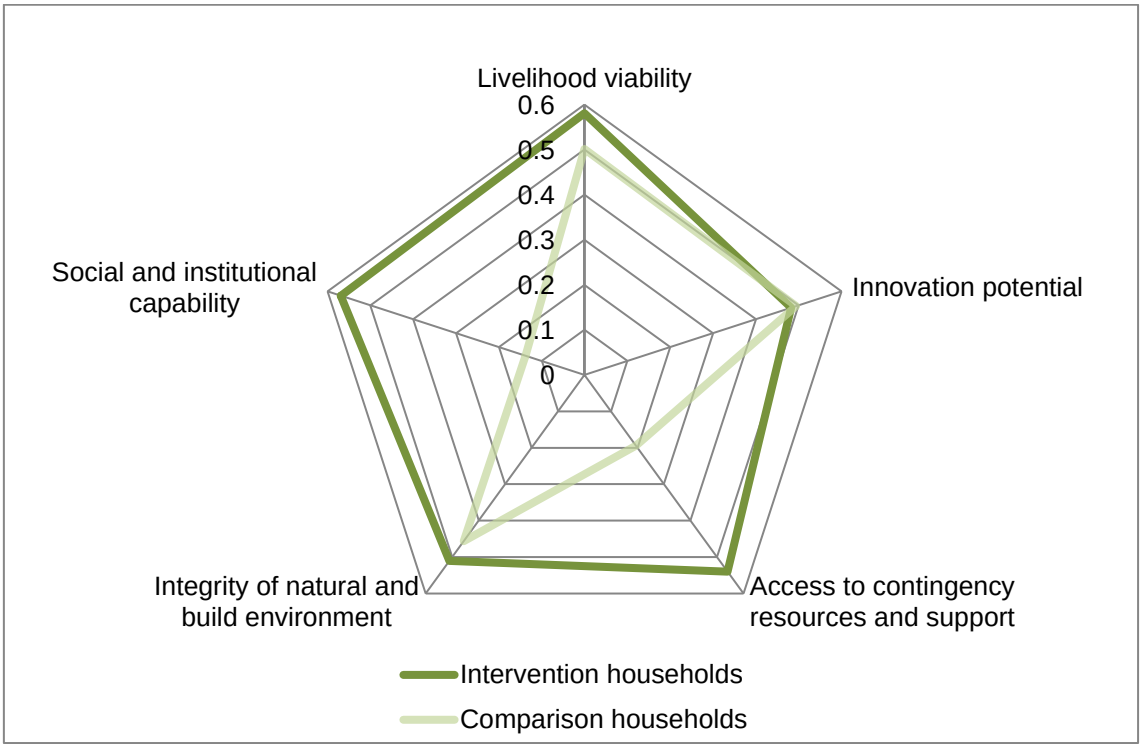
The results in Column 5, which relate to attitudes about women's representation in the community, are less directly linked to the project activities, and should therefore increase our confidence that the project had brought about more lasting changes in social and institutional capability. Respondents were asked whether they agreed more

with (1) ‘Our community relies on men’s decisions about our protection during times of crisis’ or (2) ‘Women in this community have influenced the plan of action we take in times of crisis’. Those respondents reporting that they agreed more with the second statement were scored positively on this indicator. There was a more than three-fold difference between the proportion of project and non-project households, reporting that women were represented in the community during times of crisis. We should emphasise that the sex of the household head and the sex of the respondent interviewed were not statistically significantly different between the intervention and comparison groups in the matched sample. We should therefore be confident that this result is not simply down to baseline differences between project and non-project households.

6.9.6 Comparing across dimensions

To provide an overall picture of the project’s impact across the five dimensions of resilience considered for this Effectiveness Review, we present the evidence for each dimension-specific index. These indices are created by average across all the indicators for each dimension. The results, using the matched sample, are shown in Figure 6.2.

Figure 6.2: Dimension-specific resilience indices



This diagram reinforces the main messages we presented in Tables 6.19–6.23. The most profound impacts of the project have occurred for the dimensions of ‘access to contingency resources and support’ and ‘social and institutional capability’. There were also modest improvements in terms of ‘livelihood viability’ and the ‘integrity of the natural and built environment’. However, the project did not have a statistically significant impact on ‘innovation potential’.

7 CONCLUSIONS

7.1 CONCLUSIONS

We use this section to summarise the main findings from Section 6.

The results clearly show that the project households in our survey participated far more in the activities implemented by the project than comparison households. We can be especially confident that the project increased participation in community groups, because we controlled for baseline group participation using recalled data in our analysis. It is also clear that stockpiling of emergency resources, trainings (such as those related to first aid and WASH), and construction of small-scale flood mitigation structures had been carried out more widely in project communities. However, these activities were by no means universal across the project VDCs and neither were they absent from the comparison VDCs, most likely because of government initiatives like ODF and PAF.

Our data suggest that the project had positive effects on the rate of livestock ownership and the number of types of animals that households owned. These effects appear to have been concentrated among large animals, such as cows, buffalos, and oxen. However, even though many of the livestock-related project activities were channelled through WEGs, there was no strong evidence that women's responsibility for households' animals was affected by the project.

While the project did not affect households' decisions over whether or not to engage in farming, there were positive and significant effects with regard to the amount of land that households farmed and the number of crops in their portfolio. The project had especially large positive impacts on the cultivation of drought-resistant varieties of rice and wheat and farming crops in the off-season. Project households were also more likely to sell their agricultural produce. Vitally, women's responsibility for growing crops was found to be moderately higher among project households than those in the comparison group. Nevertheless, there were no clear effects on the use of organic fertilisers, despite the project having provided training and inputs for the use of vermi-compost.

In contrast to the results for agricultural activities, the project did not have clear positive effects on households' engagement in non-farm income-generating activities. If anything, project households were less likely to engage in non-farm livelihood strategies than comparison households, although these results were largely driven by a reduction in the proportion of households engaging in casual labour, such as construction or working short-term in a factory. The main puzzle emanating from these results is that the project did not increase the proportion of households that were operating their own businesses, especially as credit and savings constraints were, in fact, relaxed among the project households. This may be because direct support for new types of livelihood activities was not built into the project until its later stages. However, this may also reflect the fact that the project was successful in supporting agricultural productivity, so non-farm income-generating activities were relatively less lucrative.

Although this Effectiveness Review was focused on resilience, we included some indicators to evaluate the impact on WASH behaviours, given their importance for the project logic. However, the project's apparent impacts on WASH were mixed. Although project households had better livestock storage practices, they relied on risky water sources – such as ponds, streams, and rivers – for their drinking more than the comparison households in our sample. However, it is important to note that these

patterns in drinking water sources were present before the start of the project, so it is very unlikely that the negative differences between the project and comparison households resulted from project activities. Rather, the project was unable to offset the pre-existing differences between project and non-project VDCs.

We included a measure of wealth in our analysis to test whether there were any differences in terms of final outcomes between project and comparison households in our sample – that is, whether project households had been able to improve their well-being in spite of shocks, stresses and uncertainty. There were, however, no clear effects of the project in terms of household wealth. In part, this may be because the survey work was carried out when households in the project areas faced only limited drought, and some localised flooding and cold waves. It is difficult to examine households' resilience using final well-being indicators when they are not currently affected, or have not recently been affected, by shocks and stresses. Also, it may take a long time for projects of this type to influence outcomes like wealth, but the survey work was carried out before the project closure.

The main outcome, on which this Effectiveness Review focused, was resilience. We created a list of indicators on the basis of discussions with project staff and focus group discussions with local communities, which could plausibly have influenced households' ability to deal with shocks, stresses, and uncertainty in the future. These were structured under the five dimensions developed for Oxfam GB's approach for measuring resilience. The indicators were also aggregated to produce a set of overall resilience indexes. We experimented with different strategies for weighting the indicators and dimensions to best reflect resilience in the local context.

Regardless of the weighting method used, the project had substantial positive effects on the resilience index. The base resilience index was a little under 50 percent higher in the project households than the non-project households in the sample. However, the largest impacts of the project were restricted to the dimensions of 'access to contingency resources and support' and 'social and institutional capability'. These were the areas on which the project worked most directly, and therefore where the majority of the indicators were most closely related to the project activities. Put differently, many of these indicators were 'output-related'. Therefore, we should be somewhat cautious about interpreting the overall increase in the resilience index as necessarily signalling changes in some of the more profound components of resilience that were high up the project's logic model or even outside the project logic.

However, the project brought about positive increases in some resilience indicators that were not directly related to the project activities. Some of the positive effects on livestock rearing and crop cultivation practice can be understood directly in terms of resilience. Moreover, dietary diversity, access to credit, and the proportion of households taking action to protect local natural resources (like forests) were all higher in the project households. These types of indicators can be regarded as relatively high-level outcomes. We therefore believe it is important to consider each resilience indicator carefully, rather than relying on the overall resilience indices, when interpreting the results from this Effectiveness Review.

7.2 PROGRAMME LEARNING CONSIDERATIONS

Take a more holistic approach to evaluation design, including discussions around suitable indicators and the possible establishment of a comparison group, at the start of future projects.

The quasi-experimental methodology deployed in this Effectiveness Review was successful due to the extensive inputs of project and partner staff during the fieldwork. The discussions around suitable comparison VDCs and the selection of good indicators of resilience were particularly rigorous and inclusive. In part, this was because the set-up for the Effectiveness Review coincided with a monitoring visit, so it was possible to involve a number of representatives from all of the partner organisations in the process. However, it would have been better to have these discussions before the project activities began. As a minimum, this would have helped create a more comprehensive Monitoring, Evaluation, and Learning plan with a wide range of resilience indicators. Going further, it would also have been possible to implement a more robust evaluation design if the comparison group had been established before the project started and baseline data had been collected in project and non-project VDCs. To support this process in the future, the impact evaluation team should provide tools and resources to help project and programme staff with quasi-experimental evaluation designs, especially on the subject of setting up a comparison group. Therefore, this learning point applies not only for programme staff, but also those working on monitoring and evaluation.

Build the support for new livelihood activities into the project at its inception rather than during the later stages of its implementation.

The results suggest that the project had far stronger effects on indicators associated with disaster risk reduction than livelihoods. In particular, there were no clear positive effects on project households' likelihood of engaging in non-farm income-generating activities, and there was only limited evidence that project households had a greater propensity to cultivate crops with new technologies, such as vermi-compost. One possible explanation for this is that many of the activities designed to support new livelihoods were only implemented in the latter half of the project's lifetime. Integrating these activities into future projects from the inception phase may increase the chances of effecting greater change in households' livelihood strategies.

Focus more directly on promoting project households' access to markets to encourage further innovation and adaptation.

Overall, there was limited evidence that the project generated clear positive effects on households' potential to innovate. This was in spite of project households reporting that their access to credit was improved. This may be because project households had, if anything, less access to markets where they could sell their produce than the comparison households in our sample. The cooperatives set up by the project could potentially remedy this issue, so they should receive special attention to ensure households can sell their crops or off-farm goods. It may also be possible to link these cooperatives to formal banking institutions, to further relax the credit constraints faced by households and, potentially, make the borrowing required to facilitate new adaptations even less costly.

Consider avenues for scaling-up the project's successes around local disaster risk reduction plans to other communities.

The most sizeable changes that were generated by the project related to households' awareness of and confidence in disaster management plans at community, VDC and district levels. However, in the non-project communities, awareness of the content of these plans – even district-level plans – was strikingly low. Despite not being the focus of this evaluation, these results suggest that the district level-activities of the project, such as supporting the planning activities of district governments, were only affecting project communities and not the non-project communities in the sample. It is important to find ways of building awareness in VDCs across the district as a whole to ensure these types of district- and national-level advocacy activities are successful.

APPENDIX 1: THRESHOLDS FOR CHARACTERISTICS OF RESILIENCE

Dimension	Characteristic	Threshold: A household scores positively if...	Directly connected to project logic?
Livelihood viability	Ownership of productive assets	Household owns five or more small assets, or two or more large assets, or four small assets and one large asset. (Small assets: ploughs, hoes, spades, threshers, scythes, carts. Large assets: grinding mills, generators, motorbikes, solar panels, computers, tractors, sewing machines.)	Yes
	Ownership of land	Household owns any land, with legal document (*lal purja*).	No
	Ownership of livestock	Household owns at least one large livestock animal (ox, cow, buffalo) and has at least two different types of animal.	Yes
	Livelihood diversification	Household engaged in any non-farm activity in the past 12 months.	Yes
	Crop diversification	Household grew three or more types of crop in the past 12 months.	Yes
	Cultivation of drought-resistant crops	Household has cultivated any drought-resistant varieties of rice, wheat, and/or maize, in the past 12 months.	Yes
	Dietary diversity.	Household members consumed a carbohydrate source on each day of the past seven days, a protein source on at least three of the past seven days, and vegetables on at least three of the past seven days.	No
	WASH attitudes	Respondent agrees more with Option 2 out of: 1. "It is OK to use ash, mud, or plain water to wash your hands because soap is too expensive." 2. "It is important to use soap to wash your hands, no matter what the price."	Yes
	Livestock hygiene practices	Household does not keep livestock in the main kitchen of the household, or in an adjacent room on the same floor/storey.	No
Innovation potential	Attitude to change	Respondent agrees more with Option 1 out of: 1. "We should not be afraid to try new and different livelihood activities – sometimes they are better than the traditional livelihood activities." 2. "It is best to continue doing what we already know and do well, rather than experimenting with new approaches."	No
	Access to credit	Household would be able to borrow 10,000 rupees to invest in a business opportunity from someone other than friends/family or a local moneylender.	Yes
	Awareness of climate change	Respondent agrees more with Option 2 out of: 1. "10 to 20 years into the future, the weather patterns in this area will be similar to those in the past." 2. "The frequency and severity of droughts and floods in this area continues to increase."	No

Dimension	Characteristic	Threshold: A household scores positively if...	Directly connected to project logic?
	Adoption of innovative practices	Household has adopted any innovations that were not directly related to the project activities.	No
	Access to markets (in monsoon season)	Household had no difficulties reaching nearest market during the monsoon season.	Yes
Access to contingency resources and support	Participation in community groups	Household members participate in at least two community groups	Yes
	Social connectivity	Household could rely on friends inside and/or outside the community if they needed to borrow 10,000 rupees to invest in a business opportunity AND respondent agrees more with Option 1 out of: 1. "We would help our neighbours with food, money, or other commodities when they face hard times, and believe they would do the same for us." 2. "It is each household's own responsibility to ensure their needs are met."	Yes
	Savings	Household members could survive for 30 days on the money they had saved, if they had an emergency and had to stop working.	Yes
	Formal earnings	Household had at least one member working in regular paid formal job in the previous 12 months.	No
	Awareness of community disaster management plan	Household knows that the community has a disaster management plan and they are at least partly aware of its contents.	Yes
	Awareness of VDC disaster management plan	Household knows that the VDC has a disaster management plan and they are at least partly aware of its contents.	Yes
	Awareness of district management plan	Household knows that the district has a disaster management plan and they are at least partly aware of its contents.	Yes
	Direct access to emergency items	Household reports that at least one type of back-up resource is available in the community in times of crisis (warm clothes, blankets, megaphones/loudspeakers, first aid kits, lifejackets, tarpaulin, back-up food supplies, radios/other communications links).	Yes
Integrity of the natural and built environment	Availability of drinking water	Household gets drinking water from tap in the house, well, or deep borehole.	Yes
	Irrigation for agriculture	Household irrigates more than 50 percent of its agricultural land.	Yes
	Protection of house from flood	Household is artificially raised up.	No
	Protection of house from fire	Household has non-flammable materials for both its walls and roof.	No
	Use of organic fertiliser	Household uses organic fertiliser (either traditional fertiliser or vermi-compost).	Yes
	Caring for forest/greenery	Household participates in a community forest group AND/OR respondent agrees more with Option 1 out of: 1. "Sometimes, members of the community must reduce their use of natural resources, like wood from the forest, to make sure we have those resources in the future."	No

Dimension	Characteristic	Threshold: A household scores positively if...	Directly connected to project logic?
		2. "Natural resources, like wood from the forest, are important for our livelihoods, so we should be able to access them whenever we want."	
Social and institutional capability	Participation in disaster planning process	Household participated in the community's disaster planning process in the previous two years.	Yes
	Early-warning system	Household is at least 'somewhat confident' that they will receive early warning information about the coming of flooding in a reliable and timely manner, and that they will be able to take appropriate action.	Yes
	Awareness of local leaders' planning activities	Respondent knows that local leaders/disaster management committees are making plans to enable the community to withstand flooding, fire, earthquake, drought and/or cold wave.	Yes
	Women's representation in the community	Respondent agrees more with Option 1 out of: 1. "The needs of women are well represented to VDC-level authorities." 2. "The VDC-level authorities mainly take into account the needs of men."	Yes
	Registration of community groups	Household participates in at least one community group that is formally/legally registered with government authorities.	Yes

APPENDIX 2: BASELINE STATISTICS BEFORE MATCHING

Table A2.1: Descriptive statistics before matching

	Intervention mean	Comparison mean	Difference	Standard error
Household size	6.36	6.13	0.23	0.21
Proportion of household members who are children (less than 15 years) (%)	33.66	32.01	1.65	1.51
Proportion of household members who are school age (7 to 18 years) (%)	25.43	23.20	2.22	1.42
Proportion of household members who are elderly (more than 65 years) (%)	5.27	6.36	-1.09	1.08
Proportion of household members who are male (%)	34.82	35.49	-0.67	1.17
Proportion of household members who have completed primary educ. (%)	28.96	23.04	5.93**	1.84
Proportion of household members who have completed secondary educ. (%)	11.75	8.21	3.54**	1.32
Proportion of household members who were present at the time of the survey (%)	91.85	89.62	2.23*	1.09
Proportion of households that existed in 2010	97.50	96.98	0.52	1.23
Household head is male (%)	77.86	80.15	-2.29	2.99
Age of household head (years)	47.19	47.04	0.15	1.01
Household head can read and write a simple letter (%)	46.43	29.30	17.13***	3.48
Household head completed primary education (%)	28.21	19.47	8.74**	3.07
Household head completed secondary education (%)	14.64	8.32	6.33**	2.26
Household head is Janajati (%)	12.50	16.64	-4.14	2.65
Household head is Dalit (%)	26.79	45.94	-19.15***	3.55
Household head is Brahmin/Chetri (%)	10.36	4.16	6.20***	1.79
Household head is Madeshi (%)	31.07	18.53	12.55***	3.08
Number of minutes it took to walk to the nearest road in 2010	29.96	28.72	1.24	2.39
Number of minutes it took to walk to the nearest market in 2010	72.61	50.86	21.76***	4.50
Farming crops (%)	72.50	49.72	22.78***	3.57
Casual labour (e.g. construction, carpentry, masonry) (%)	37.50	50.09	-12.59***	3.66
Household business (e.g. retail shop, hotel, tea shop) (%)	7.50	2.65	4.85**	1.50
Regular, paid employment (e.g. as a teacher or nurse) (%)	11.43	7.37	4.06	2.09
Household was in the lowest 20% of the wealth distribution, in 2010(%)	19.29	20.42	-1.13	2.96
Household was in the second 20% of the wealth distribution, in 2010 (%)	17.50	21.36	-3.86	2.96
Household was in the third 20% of the wealth distribution, in 2010 (%)	17.14	21.55	-4.41	2.96
Household was in the fourth 20% of the wealth distribution, in 2010 (%)	21.79	19.09	2.69	2.96

Household was in the highest 20% of the wealth distribution, in 2010 (%)	24.29	17.58	6.71*	2.94
Household participated in any community groups, in 2010 (%)	55.00	24.57	30.43***	3.37
Number of community groups in which household participated, in 2010 (%)	1.64	0.41	1.23***	0.12
Household participated in Women's Empowerment Group(s), in 2010 (%)	22.50	4.91	17.59***	2.23
Observations	280	529	809	

The construction of the wealth index is described in Section 6
Variables dated 2010 are estimates, based on recall data
* p < 0.1, ** p < 0.05, *** p < 0.01

APPENDIX 3: METHODOLOGY USED FOR PROPENSITY SCORE MATCHING

The results presented in Section 6 of this report were estimated using propensity-score matching (PSM). PSM is a statistical technique that allows us to estimate the effect of an intervention by accounting for the covariates that predict receiving the intervention, or 'treatment'. The idea behind PSM is to match similar individuals in the treatment or intervention group to those in the control or comparison group, based on observed characteristics at baseline. After each participant is matched with a non-participant, the average treatment effect on the treated (those who benefited from the intervention) is equal to the difference in average outcomes of the intervention and the comparison groups after project completion. This appendix describes and tests the specific matching procedure employed in this Effectiveness Review. A practical guide on the different approaches to matching may be found in Caliendo and Kopeinig (2008).

Estimating propensity scores

Finding an exact match for treated individuals, based on various baseline characteristics would be very hard to implement in practice. Rosenbaum and Rubin (1983) demonstrated that a 'propensity score' could summarise all this information in one single variable. The propensity score is defined as the conditional probability of receiving the intervention given background variables. Specifically, propensity scores are calculated using a statistical probability model (e.g. probit or logit) to estimate the probability of participating in the project, conditional on a set of characteristics.

Table A3.1 and shows the variables used to estimate the propensity score. Here, we report the marginal effects at the mean, and the corresponding standard errors. Following Caliendo and Kopeinig (2008), only variables that influence the participation decision, but which are not affected by participation in the project, were included in our matching model. In the table, the dependent variable corresponds to whether or not an individual received the intervention – it is equal to 1 if the household belongs to one of the communities that benefited from the project activities, and 0 otherwise. The coefficients in the table correspond to the marginal effects, i.e. the change in the probability of receiving the intervention if the independent variable is increased by one.

Defining the region of common support

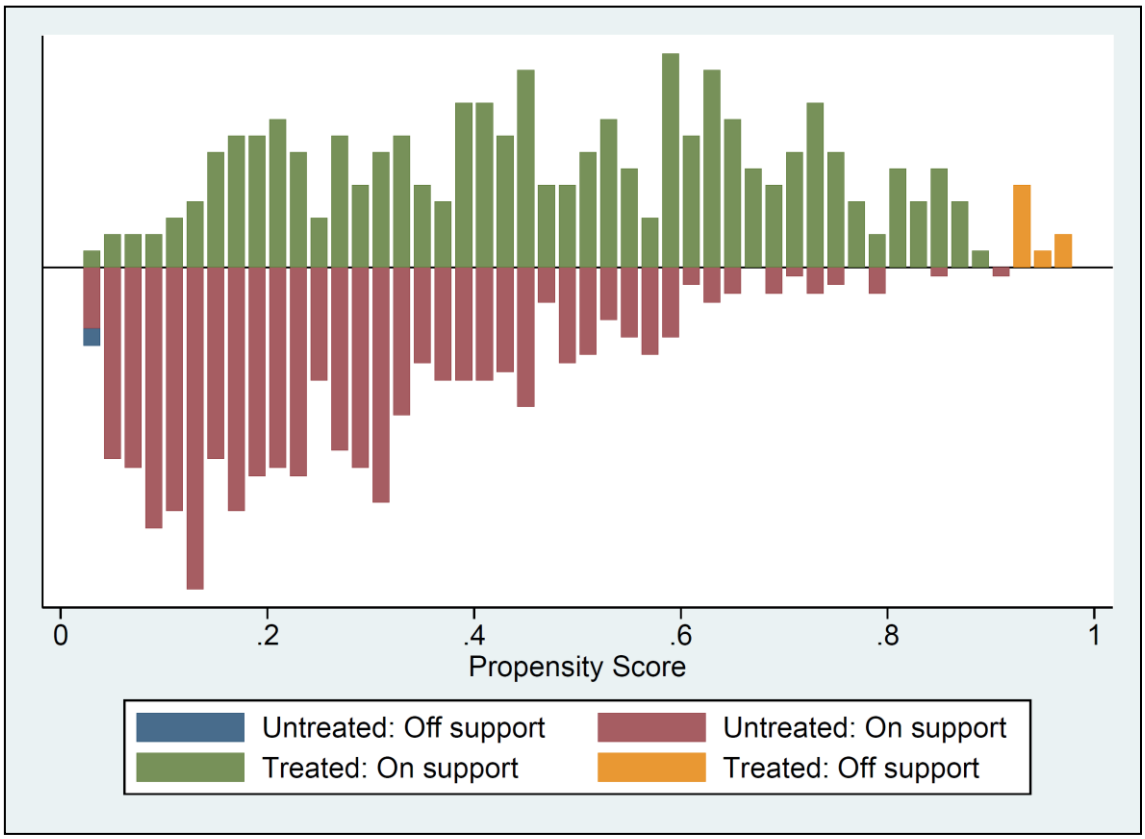
After estimating the propensity scores, we need to verify that there is a potential match for the observations in the intervention group with those from the comparison group. This means checking that there is *common support*. The area of common support is the region where the propensity score distributions of the intervention and comparison groups overlap. The common support assumption ensures that each 'treatment [intervention] observation has a comparison observation "nearby" in the propensity score distribution' (Heckman, LaLonde and Smith, 1999). Figure A3.1 shows the propensity score density plots for both groups. We observe that, although the distributions of propensity scores are clearly different between the intervention and comparison groups in each case, there is a reasonably good area of overlap between the groups. However, in constructing the model for household-level outcomes, eight observations from the intervention group and two observations from the comparison group were dropped because there was not a suitable match for them.

Table A3.1: Estimating the propensity score

	Marginal effect	Standard error	p-value
Household head can read and write a simple letter? (1=Y, 0=N)	0.15***	0.04	0.00
Household head is male? (1=Y, 0=N)	-0.12*	0.05	0.01
Household head is Janajati	-0.16***	0.05	0.00
Household head is Dalit	-0.17***	0.04	0.00
Log of household size	0.07	0.04	0.13
Time taken to walk to nearest market in 2010 (minutes)	0.00***	0.00	0.00
HH undertook farming of crops in 2010? (1=Y, 0=N)	0.17***	0.04	0.00
HH engaged in any non-farm income-generating activities in 2010? (1=Y, 0=N)	-0.08*	0.04	0.04
HH had a household business in 2010 (1=Y, 0=N)	0.26**	0.10	0.01
HH was in the second 20% of the wealth distribution, in 2010	-0.14**	0.05	0.00
HH was in the third 20% of the wealth distribution, in 2010	-0.20***	0.05	0.00
HH was in the fourth 20% of the wealth distribution, in 2010	-0.18***	0.05	0.00
HH was in the highest 20% of the wealth distribution, in 2010	-0.20***	0.05	0.00
HH participated in any community groups in 2010 (1=Y, 0=N)	0.29***	0.04	0.00
Observations	809		

The construction of the wealth index is described in Section 6. Variables dated 2010 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

Figure A3.1: Histogram of propensity scores in the intervention and comparison groups



Matching intervention households to comparison households

Following Rosenbaum and Rubin (1983), households were matched on the basis of their propensity scores. The literature has developed a variety of matching procedures. After a series of checks, we decided to employ the kernel matching algorithm for the results presented in this Effectiveness Review. Kernel matching assigns more weight to the closest comparison group observations that are found within a selected 'bandwidth'. Thus 'good' matches are given greater weight than 'poor' matches. We used the *psmatch2* module in Stata with a bandwidth of 0.075 and restricted the analysis 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.[24]

Checking balance

For PSM to be valid, the intervention group and the matched comparison group need to be balanced. In other words, the intervention and comparison groups need to be similar in terms of their observed characteristics. The most straightforward method of doing this is to test whether there are any statistically significant differences in baseline covariates between both groups in the matched sample. The balance of each of the matching variables after kernel matching is shown in Table A3.2. There are no statistically significant differences between the intervention and comparison groups for any of the matching variables used in the matched sample For all of these variables, the *p*-values for the difference in means tests are larger than 0.2. We can therefore conclude in each case that we have found a satisfactory match for the *observable* variable our sample.

Table A3.2: Balancing test on matching variables

	Intervention group mean	Comparison group mean	p-value
Household head can read and write a simple letter? (1=Y, 0=N)	0.45	0.43	0.74
Household head is male? (1=Y, 0=N)	0.78	0.77	0.91
Household head is Janajati	0.13	0.1	0.41
Household head is Dalit	0.28	0.27	0.87
Household size	6.42	6.63	0.46
Time taken to walk to nearest market in 2010 (minutes)	65.67	56.84	0.2
HH undertook farming of crops in 2010? (1=Y, 0=N)	0.72	0.68	0.4
HH had any off-farm income at the start in 2010? (1=Y, 0=N)	0.47	0.48	0.92
HH had a household business in 2010 (1=Y, 0=N)	0.08	0.09	0.84
HH was in the second 20% of the wealth distribution, in 2010	0.18	0.22	0.38
HH was in the third 20% of the wealth distribution, in 2010	0.17	0.18	0.64
HH was in the fourth 20% of the wealth distribution, in 2010	0.21	0.17	0.18
HH was in the highest 20% of the wealth distribution, in 2010	0.24	0.25	0.83
HH participated in any community groups in 2010 (1=Y, 0=N)	0.54	0.56	0.56

The construction of the wealth index is described in Section 6
Variables dated 2010 are estimates, based on recall data
* p < 0.1, ** p < 0.05, *** p < 0.01

As discussed in Section 4.2, we also consider whether the sample is balanced in terms of the land holdings and the education levels of adults in the household. This enables us to check whether the screening questions we used for the comparison group were able to replicate the process that was originally used to select project participants. We measure land holdings at the start of the project using recalled data on the number of katthas of land that were owned by the household – with the *lal purja* document legal document – in 2010. We then present the proportion of adults in the household that had completed primary education, as well as the education levels of the household head. Again, there are no statistically significant differences between these variables, in the matched sample. This should reassure us that the use of different sampling procedures in the project VDCs and the non-project VDCs has not led to potential bias in our results.

Table A3.3: Balancing test on matching variables

	Intervention group mean	Comparison group mean	p-value
Number of katthas of land HH owned (with legal document – *lal purja*) in 2010	6.45	6.29	0.86
All adults in the HH have completed primary education? (1=Y, 0=N)	0.01	0.00	0.19
Prop. adults in the HH who have completed primary education	0.24	0.23	0.64
Household head has completed primary education? (1=Y, 0=N)	0.28	0.27	0.83
Household head has completed secondary education? (1=Y, 0=N)	0.14	0.13	0.82

APPENDIX 4: ROBUSTNESS CHECKS

In order to assess the robustness of the results presented in Section 6, a series of checks were carried out to determine whether the main findings of this report are sensitive to the estimation procedure – propensity score matching with the kernel method – that was used to control for observable differences between the intervention and comparison group. This appendix presents five types of robustness checks.

1 Multivariate regression

The first robustness test we run is to estimate the impact of project participation using an ordinary least squares (OLS) regression. The main idea behind OLS is to isolate the variation in the outcome variable that is due to the intervention status – the project's impact – by controlling directly for the influence that observable differences between the intervention and comparison group have on outcomes. To do this, we estimate Equation 1.[25]

Equation 1

$$Y_i = \alpha + \beta_1 \tau_i + \boldsymbol{X}_i' \boldsymbol{\delta} + \varepsilon_i$$

In Equation 1, Y_i is the dependent variable (the outcome) and $\boldsymbol{X_i}$ is a column vector of the same matching variables listed in tables A3.1 or A3.2. The intervention status is given by a dummy variable (τ_i), which takes the value of 1 if the household participated in the project and 0 otherwise. The key difference between this OLS regression model and the propensity score matching procedure used in the main report is that the OLS regression estimates a direct parametric relationship between the covariates in $\boldsymbol{X_i}$ and the dependent variable Y_i. This means it is possible to include the observations that were excluded due to being off common support in Section 6 by extrapolating the relationship between $\boldsymbol{X_i}$ and Y_i. It should be borne in mind, however, that extrapolating in this way may bias the results if the covariates are distributed very differently between the intervention and comparison group (Rubin, 2001).[26]

It is also important to note that, as with the PSM methods used in the main body of the report, OLS regressions can only account for observable differences between the intervention and comparison groups. Unobservable differences may still bias the results. In the tables that follow, only the estimate of $\beta_1$will be reported.

2 Multivariate regression with alternative matching variables

Given the importance of controlling for recalled group participation for minimising bias, we tested whether our results are sensitive to the way this information is captured by the matching variables. To do this, we altered $\boldsymbol{X_i}$, removing the dummy variable for whether or not the household participated in any community groups in 2010 and added instead: (1) the number of community groups in which the household participated in 2010, and (2) a dummy variable for whether or not the household participated in a WEG. Once again, we only report the estimates of β_1.

3 Propensity score weighting

Following the example of Hirano and Imbens (2001) we also estimate OLS regressions using exactly the same model as in Equation 1, but weighting the observations according to the propensity score. Observations are assigned weights equal to 1 for the intervention households and $\hat{P}(\boldsymbol{X_i})/(1-\hat{P}(\boldsymbol{X_i}))$ for the comparison households. The variable $\hat{P}(\boldsymbol{X_i})$ represents the probability of a household being in the intervention group, given their observable characteristics, measured through the vector of matching variables $\boldsymbol{X_i}$ – this was estimated in the probit regressions in Appendix 3. We report the estimates of β_1 in the same way as the regular OLS regressions.

4 Nearest neighbour matching

The nearest neighbour (NN) matching algorithm matches each observation from the intervention group with an observation from the comparison group that is closest in terms of their propensity score.[27] In this robustness check, we apply the NN method 'with replacement', meaning that comparison observations can be matched to intervention observations more than once.[28] In the tables below, we report the estimated differences between the intervention and comparison groups.

5 Nearest neighbour with exact district matching

When the NN matching algorithm is used it is possible to apply certain restrictions to what matches are permitted. One possibility with the data in this Effectiveness Review is to constrain the matching process so that only households within the same district are matched to each other. This eliminates the possibility, for example, that a household from Sarlahi is matched to a household from Saptari. Again, the estimated differences between the intervention and comparison group are reported in the tables below.

In the remainder of this appendix, we report these robustness checks for the main results of the report.

Table A4.1: Overall resilience indexes

	1	2	3	4
	Base resilience index	Equal dimensions resilience index	Personal dimensions resilience index	Sample dimensions resilience index
OLS regression	0.18***	0.18***	0.20***	0.21***
	(0.01)	(0.01)	(0.01)	(0.01)
N	809	809	809	809
OLS regression with alternative matching variables	0.18***	0.19***	0.20***	0.21***
	(0.01)	(0.01)	(0.01)	(0.01)
N	809	809	809	809
OLS with PS weighting	0.18***	0.18***	0.20***	0.20***
	(0.01)	(0.01)	(0.01)	(0.01)
N	799	799	799	799
Nearest neighbour	0.18***	0.17***	0.19***	0.20***
	(0.02)	(0.02)	(0.02)	(0.02)
N	799	799	799	799
Nearest neighbour with exact district matching	0.16***	0.16***	0.18***	0.18***
	(0.03)	(0.02)	(0.03)	(0.03)
N	799	799	799	799

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

Table A4.2: Indicators of livelihood viability

Part A

	1	2	3	4	5
	Ownership of productive assets	Ownership of land	Ownership of livestock	Livelihood diversification	Crop diversification
OLS regression	0.03	0.07*	0.17***	-0.02	0.09**
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
N	809	809	809	809	809
OLS regression with alternative matching variables	0.02	0.07**	0.16***	-0.01	0.11***
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
N	809	809	809	809	809
OLS with PS weighting	0.01	0.06*	0.20***	-0.07*	0.14***
	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
N	799	799	799	799	799
Nearest neighbour	-0.00	0.06	0.23***	-0.07	0.17***
	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
N	799	799	799	799	799
Nearest neighbour with exact district matching	-0.11	0.10**	0.04	-0.10**	0.13
	(0.08)	(0.05)	(0.09)	(0.05)	(0.08)
N	799	799	799	799	799

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

Part B

	6	7	8	9
	Drought-resistant crop varieties	Dietary diversity	WASH attitudes	Livestock storage practices
OLS regression	0.12***	0.13***	0.09*	0.08**
	(0.03)	(0.04)	(0.04)	(0.03)
N	809	809	809	809
OLS regression with alternative matching variables	0.10***	0.14***	0.10**	0.08***
	(0.04)	(0.04)	(0.04)	(0.03)
N	809	809	809	809
OLS with PS weighting	0.16***	0.15***	0.05	0.11***
	(0.04)	(0.04)	(0.04)	(0.04)
N	799	799	799	799
Nearest neighbour	0.11**	0.15**	0.03	0.13**
	(0.05)	(0.06)	(0.05)	(0.05)
N	799	799	799	799
Nearest neighbour with exact district matching	0.01	-0.03	-0.01	0.15*
	(0.08)	(0.09)	(0.09)	(0.08)
N	799	799	799	799

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

Table A4.3: Indicators of innovation potential

	1	2	3	4	5
	Attitudes to change	Access to credit	Awareness of climate change	Adoption of innovative practices	Access to markets
OLS regression	-0.01	0.16***	0.02	0.02	-0.01
	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
N	809	809	809	809	809
OLS regression with alternative matching variables	-0.00	0.17***	0.03	0.02	-0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
N	809	809	809	809	809
OLS with PS weighting	-0.02	0.10***	-0.03	-0.02	-0.02
	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
N	799	799	799	799	799
Nearest neighbour	-0.06	0.08*	-0.03	-0.04	-0.06
	(0.06)	(0.04)	(0.05)	(0.06)	(0.06)
N	799	799	799	799	799
Nearest neighbour with exact district matching	-0.03	0.07	-0.00	-0.02	-0.03
	(0.09)	(0.04)	(0.05)	(0.09)	(0.09)
N	799	799	799	799	799

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

Table A4.4: Indicators of access to contingency resources and support

Part A

	1	2	3	4
	Group participation	Social connectivity	Savings	Formal earnings
OLS regression	0.45***	0.22***	0.06	0.04
	(0.03)	(0.04)	(0.04)	(0.03)
N	809	809	809	809
OLS regression with alternative matching variables	0.46***	0.21***	0.05	0.02
	(0.04)	(0.04)	(0.04)	(0.03)
N	809	809	809	809
OLS with PS weighting	0.38***	0.19***	0.08*	0.02
	(0.04)	(0.04)	(0.04)	(0.03)
N	799	799	799	799
Nearest neighbour	0.32***	0.14**	0.04	0.03
	(0.05)	(0.06)	(0.07)	(0.05)
N	799	799	799	799
Nearest neighbour with exact district matching	0.38***	0.14*	0.06	0.01
	(0.08)	(0.08)	(0.05)	(0.05)
N	799	799	799	799

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

Part B

	5	6	7	8
	Awareness of community disaster management plan	Awareness of VDC disaster management plan	Awareness of district management plan	Direct access to emergency items
OLS regression	0.55***	0.52***	0.42***	0.49***
	(0.03)	(0.03)	(0.03)	(0.04)
N	809	809	809	809
OLS regression with alternative matching variables	0.54***	0.51***	0.40***	0.52***
	(0.03)	(0.03)	(0.03)	(0.04)
N	809	809	809	809
OLS with PS weighting	0.58***	0.55***	0.46***	0.49***
	(0.03)	(0.03)	(0.03)	(0.04)
N	799	799	799	799
Nearest neighbour	0.61***	0.57***	0.49***	0.53***
	(0.04)	(0.05)	(0.05)	(0.06)
N	799	799	799	799
Nearest neighbour with exact district matching	0.61***	0.57***	0.49***	0.54***
	(0.03)	(0.03)	(0.03)	(0.05)
N	799	799	799	799

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

Table A4.5 Indicators for the integrity of the natural and built environment

	1	2	3	4	5	6
	Availability of drinking water	Irrigation for agriculture	Protection of house from flood	Protection of house from fire	Use of organic fertiliser	Caring for forest/ greenery
	-0.04**	0.03	0.21***	-0.04	-0.00	0.11**
OLS regression	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
	809	809	809	809	809	809
N	-0.05***	0.01	0.19***	-0.02	0.04	0.10**
OLS regression with alternative matching variables	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
	809	809	809	809	809	809
N	-0.04**	0.07*	0.24***	-0.06	-0.03	0.11***
OLS with PS weighting	(0.02)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
	799	799	799	799	799	799
N	-0.04	0.07	0.23***	-0.06	-0.03	0.07
Nearest neighbour	(0.04)	(0.06)	(0.04)	(0.06)	(0.06)	(0.06)
	799	799	799	799	799	799
N	-0.06**	-0.09	0.25***	0.05	-0.09	0.07
Nearest neighbour with exact district matching	(0.03)	(0.08)	(0.04)	(0.04)	(0.08)	(0.08)
	799	799	799	799	799	799
N	-0.04**	0.03	0.21***	-0.04	-0.00	0.11**

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

Table A4.6: Indicators of social and institutional capability

	1	2	3	4	5
	Participation in disaster planning process	Early-warning system	Awareness of local leaders planning activities	Registration of community groups	Women's representation in the community
OLS regression	0.44***	0.56***	0.52***	0.30***	0.27***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
N	809	809	809	809	809
OLS regression with alternative matching variables	0.45***	0.56***	0.51***	0.31***	0.25***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)
N	809	809	809	809	809
OLS with PS weighting	0.44***	0.55***	0.54***	0.31***	0.31***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)
N	799	799	799	799	799
Nearest neighbour	0.43***	0.58***	0.54***	0.29***	0.31***
	(0.03)	(0.05)	(0.04)	(0.06)	(0.04)
N	799	799	799	799	799
Nearest neighbour with exact district matching	0.44***	0.58***	0.57***	0.21**	0.27***
	(0.03)	(0.04)	(0.03)	(0.08)	(0.05)
N	799	799	799	799	799

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

Table A4.7: Wealth

	1	2
	Normalised wealth index	Difference in normalised wealth index
OLS regression	0.14*	0.15*
	(0.06)	(0.06)
N	809	809
OLS regression with alternative matching variables	0.21***	0.21***
	(0.06)	(0.06)
N	809	809
OLS with PS weighting	0.06	0.08
	(0.08)	(0.07)
N	799	799
Nearest neighbour	-0.01	0.03
	(0.16)	(0.11)
N	799	799
Nearest neighbour with exact district matching	0.04	0.16**
	(0.13)	(0.07)
N	799	799

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

Table A4.8: Number of animals owned at the time of the survey

	1	2	3	4	5	6
	Cows	Buffalo	Oxen	Goats	Pigs	Poultry
OLS regression	0.25*	0.13	0.10	0.33*	0.25*	0.13
	(0.10)	(0.07)	(0.06)	(0.16)	(0.10)	(0.07)
N	809	809	809	809	809	809
OLS regression with alternative matching variables	0.23**	0.11	0.11*	0.30*	0.23**	0.11
	(0.10)	(0.07)	(0.06)	(0.16)	(0.10)	(0.07)
N	809	809	809	809	809	809
OLS with PS weighting	0.36***	0.21**	0.17***	0.26	0.36***	0.21**
	(0.13)	(0.10)	(0.06)	(0.19)	(0.13)	(0.10)
N	799	799	799	799	799	799
Nearest neighbour	0.41***	0.14	0.14*	0.35	0.41***	0.14
	(0.15)	(0.11)	(0.07)	(0.29)	(0.15)	(0.11)
N	799	799	799	799	799	799
Nearest neighbour with exact district matching	0.07	0.18	-0.03	-0.13	0.07	0.18
	(0.19)	(0.12)	(0.14)	(0.32)	(0.19)	(0.12)
N	799	799	799	799	799	799

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

Table A4.9: Overall engagement in agricultural activities

	1	2	3	4
	Household grew any crops in the past year	Number of katthas of land household cultivated	Number of types of crops the household grew in the past year	Number of types of non-cereal crops the household grew in the past year
OLS regression	0.03	1.94**	0.45***	0.44***
	(0.02)	(0.74)	(0.11)	(0.09)
N	809	809	809	809
OLS regression with alternative matching variables	0.03	1.97***	0.52***	0.48***
	(0.02)	(0.75)	(0.11)	(0.10)
N	809	809	809	809
OLS with PS weighting	0.03	2.40***	0.60***	0.58***
	(0.02)	(0.74)	(0.12)	(0.10)
N	799	799	799	799
Nearest neighbour	0.07	1.93*	0.75***	0.62***
	(0.06)	(1.12)	(0.22)	(0.16)
N	799	799	799	799
Nearest neighbour with exact district matching	-0.01	-0.16	0.50*	0.54***
	(0.08)	(1.95)	(0.28)	(0.18)
N	799	799	799	799

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

Table A4.10: Off-season crop growing and drought-resistant varieties

	1	2	3	4
	Household grew any crops in the off-season	Household cultivated drought-resistant rice	Household cultivated drought-resistant wheat	Household cultivated drought-resistant maize
OLS regression	0.11***	0.13***	0.11**	0.00
	(0.03)	(0.03)	(0.03)	(0.02)
N	809	809	809	809
OLS regression with alternative matching variables	0.09***	0.12***	0.09***	0.01
	(0.03)	(0.03)	(0.03)	(0.02)
N	809	809	809	809
OLS with PS weighting	0.14***	0.17***	0.14***	0.01
	(0.03)	(0.04)	(0.04)	(0.02)
N	799	799	799	799
Nearest neighbour	0.12**	0.14***	0.08*	0.00
	(0.05)	(0.05)	(0.05)	(0.02)
N	799	799	799	799
Nearest neighbour with exact district matching	0.12***	0.02	-0.01	0.00
	(0.04)	(0.08)	(0.08)	(0.02)
N	799	799	799	799

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

Table A4.11: Non-farm livelihood activities

	1	2	3	4	5	6
	Household engaged in any non-farm income-generating activities	Number of non-farm income-generating activities that the household engaged in	Female household member engaged in any non-farm income-generating activities	Casual labour (e.g. construction, working short-term in a factory)	Household business (e.g. retail shop, hotel, tea shop)	Regular, paid formal job (e.g. teacher, nurse, public sector, long-term factory
OLS regression	-0.02	0.01	0.02	-0.06	0.03	0.04
	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)
N	809	809	809	809	809	809
OLS regression with alternative matching variables	-0.01	0.01	0.01	-0.04	0.03	0.02
	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)
N	809	809	809	809	809	809
OLS with PS weighting	-0.07*	-0.05	-0.01	-0.08**	0.01	0.02
	(0.03)	(0.05)	(0.03)	(0.04)	(0.03)	(0.03)
N	799	799	799	799	799	799
Nearest neighbour	-0.07	-0.04	-0.04	-0.06	-0.02	0.03
	(0.05)	(0.08)	(0.05)	(0.06)	(0.05)	(0.05)
N	799	799	799	799	799	799
Nearest neighbour with exact district matching	-0.10**	-0.09	0.02	-0.13**	0.02	0.01
	(0.05)	(0.10)	(0.07)	(0.05)	(0.08)	(0.05)
N	799	799	799	799	799	799

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

Table A4.12: Livestock storage practices

	1	2	3
	Same room	In different structure	Livestock kept outside with no shed
OLS regression	-0.07*	0.08*	-0.05*
	(0.03)	(0.04)	(0.02)
N	608	608	608
OLS regression with alternative matching variables	-0.05	0.08**	-0.04**
	(0.03)	(0.04)	(0.02)
N	608	608	608
OLS with PS weighting	-0.13***	0.13***	-0.07**
	(0.04)	(0.04)	(0.03)
N	600	600	600
Nearest neighbour	-0.16**	0.15**	-0.08*
	(0.06)	(0.06)	(0.04)
N	600	600	600
Nearest neighbour with exact district matching	-0.01	0.15*	-0.04
	(0.04)	(0.09)	(0.04)
N	600	600	600

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

Table A4.13: Main source of drinking water for the household

	1	2	3
	A tap in the house	Well/borehole	Other (e.g. pond, river, stream)
OLS regression	-0.04	-0.01	0.04**
	(0.02)	(0.02)	(0.01)
N	809	809	809
OLS regression with alternative matching variables	-0.06***	0.01	0.05***
	(0.02)	(0.02)	(0.01)
N	809	809	809
OLS with PS weighting	-0.03	-0.01	0.04**
	(0.03)	(0.03)	(0.02)
N	799	799	799
Nearest neighbour	0.03	-0.06	0.04
	(0.03)	(0.05)	(0.04)
N	799	799	799
Nearest neighbour with exact district matching	-0.03	-0.03	0.06**
	(0.04)	(0.04)	(0.03)
N	799	799	799

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

In general, these robustness checks support the main results. Even if the coefficient is no longer statistically significant, there are only a few instances when the sign on the coefficient differs from what was presented in Section 6, and these cases do not apply to the headline results. The nearest neighbour matching algorithm seems to cause the most change to the results, especially when exact district matching is imposed.

APPENDIX 5: SUBGROUP ANALYSIS

In this appendix, we consider whether project had differential effects depending on the gender of the household head. As the descriptive statistics in Appendix 3 show, the ratio of male-headed households to female-headed households was approximately 4:1.

To assess whether the gender of the household head alters the projects' impact, we run an OLS regression, similar to the robustness checks in Appendix 4. However, we add a so-called 'interaction' variable, to the equation, which is simply the intervention status (τ_i) multiplied by a dummy variable for the gender of the household head (D_i). This variable takes the value 1 if the household head was male and 0 if the household head was female. The regression model also includes the matching variables as covariates ($\boldsymbol{X}_i$), to control for observable baseline differences between the project and non-project households when estimating the effects of the project. The regression equation estimated is shown in Equation 2.[29]

Equation 2

$$\boldsymbol{Y_i = \alpha + \beta_1 \tau_i + \beta_2 (\tau_i \times D_i) + \beta_3 D_i + X_i' \delta + \varepsilon_i}$$

If the coefficient β_2 is statistically significant, this suggests that there have been differential effects on male- and female-headed households.

We focus this subgroup analysis on the main resilience indexes. The results are shown in Table A5.1.

Table A5.1: Differential effects on resilience by gender of the household head

	1	2	3	4
	Base resilience index	Equal dimensions resilience index	Personal dimensions resilience index	Sample dimensions resilience index
Intervention	0.16***	0.16***	0.18***	0.19***
	(0.02)	(0.02)	(0.02)	(0.02)
Intervention * Gender of Household Head	0.03	0.03	0.02	0.02
	(0.02)	(0.02)	(0.03)	(0.02)
Gender of Household Head (1=Male, 0=Female)	-0.02*	-0.02*	-0.02*	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	809	809	809	809

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

The results in Table A5.1 do not suggest that the project had differential effects on male- and female-headed households because none of the interaction terms are statistically significant. However, there is some tentative evidence that, on average, resilience was moderately lower in male-headed households than female-headed households. This difference – of around 2 percentage points – is significant at the 10 percent level in Columns 1, 2, and 3.

APPENDIX 6: WEIGHTING EXERCISE

During the questionnaire, respondents were asked to provide their perceptions of which dimensions of resilience that they thought were most important. To do this, the enumerators conducted a so-called 'budget allocation game' with each respondent. This involved showing the respondent a special laminated sheet, on which five pictures were drawn, each representing one of the five dimensions of resilience (see below). Next to each picture is a short description, explaining what each picture represents. The enumerators read out these descriptions in the local language to help the respondents understand what the pictured showed. The content of each picture was developed through consultation with Oxfam staff and a local artist in Ethiopia, during the fieldwork for a different Effectiveness Review. The pictures and descriptions were kept the same for the Nepal Effectiveness Review because staff both from Oxfam and from the local partner organisations suggested that they would capture the same concepts in the Nepalese context too. It was also thought that attempting to change the pictures posed a greater risk to the comparability between the two Effectiveness Reviews. The pictures were not developed to provide a 'complete' picture of each of the five dimensions. Rather they were designed to capture the main feature(s) of a particular dimension in as simple and comprehensible way as possible.

After being shown the pictures and having heard the descriptions, respondents were asked about how important they thought each dimension was for resilience. Specifically, they were asked: 'How important are the following things to you in making sure your household members have everything they need, even in difficult times?'. To show this, respondents were given 15 stones, of approximately equal sizes, and asked to place the stones on each category to show what they thought was most important for resilience. The stones were chosen to be of roughly equal size to prevent confusion. Objects with more standard sizes – such as grains, buttons, or counters – were not thought to be suitable for the context of Nepal.

The enumerator then recorded the number of stones on each dimension of resilience in the mobile devices. This would then automatically check that the number of stones added up to 15, to ensure data quality.

Figure A6.1: Sheet shown to respondents during the 'budget allocation game'

BIBLIOGRAPHY

A. Abadie and G. W. Imbens, On the failure of the bootstrap for matching estimators, *Econometric*, vol. 76 (2008), pages 1537–1557.

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

Martin Bland and Douglas G. Altman, Statistics notes: Cronbach's alpha, *British Medical Journal*, (1997).

Deon Filmer and Lant H. Pritchett, Estimating Wealth Effects without Expenditure Data – Or Tears: An Application to Educational Enrollments in States of India, *Demography*, vol. 38(1) (2001), pages 115–132.

James J. Heckman, Robert J. LaLonde and Jeffrey A. Smith, The Economics and Econometrics of Active Labor Market Programs, *Handbook of Labor Economics*, vol. 3, part A (1999), pages 1865–2097.

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

Paul R. Rosenbaum and Donald B. Rubin, The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika*, vol. 70(1) (1983), pages 41–55.

Donald B. Rubin, Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation, *Health Services & Outcomes Research Methodology*, vol. 2(3) (2001), pages 169-188.

NOTES

1 There were, on average, six VDCs participating in the project in each district.

2 Equal numbers of households were targeted for interview in each VDC for simplicity – VDCs were broadly similar in terms of their populations.

3 Sometimes this protocol for finding replacements was applied, but then the originally sampled household became available for interview. This explains why the sample sizes for some project VDCs exceed 35 and for some comparison VDCs exceed 44.

4 In only four of the project households did all the adult members report having completed primary education, Grades 1 to 5.

5 This arises due to 'classical measurement error', which attenuates effect sizes – including for basic t-tests – towards zero.

6 This approach is described in 'A Multidimensional Approach to Measuring Resilience', Oxfam GB working paper, August 2013: http://policy-practice.oxfam.org.uk/publications/a-multidimensional-approach-to-measuring-resilience-302641.

7 Oxfam International's recent programme guidelines have conceptualised resilience in terms of three capacities rather than five dimensions. The five dimensions used in this Effectiveness Review represent a specific Oxfam GB methodology for measuring resilience. Future Oxfam impact evaluations will make use of the three capacities framework instead.

8 Specifically, focus group discussions were carried out in Pattharkot VDC, Sarlahi district.

9 It should be noted that, if the dimensions are first weighted equally, each additional indicator added to a dimension is effectively down-weighted in the resulting resilience index.

10 The regression models shown in Appendix 4 do not restrict the sample to the area of common support.

11 We believe that the question about WEGs was understood more broadly in terms of other types of women's groups in the comparison group.

12 There was also a slight difference in the proportion of households owning livestock according to the recalled baseline data. We do not, therefore, attribute the difference shown in Column 1 of Table 6.5 to the effects of the project.

13 One kattha is equal to 338.63 m^2 (around 8 percent of an acre).

14 Across the sample as a whole, women were more frequently responsible for staple crops – such as rice, wheat, and potatoes – compared with other vegetables.

15 This result is not statistically significant for any of the robustness checks reported in Appendix 5.

16 Specifically the question asked: 'How far is/was the main livestock shed from the main kitchen in your household?' 0 = household has no livestock; 1 = same room; 2 = on a different floor/storey; 3 = same floor/storey and in adjacent rooms; 4 = same floor/storey and same structure but not in adjacent rooms; 5 = in different structures/huts; 6 = livestock kept outside with no shed

17 We ensure the item-rest correlation for each asset is greater than 0.1. We also ensure that Cronbach's alpha is at least 0.7, following the BMJ guidance note (Bland and Altman, 1997). The resulting list of assets included in the wealth index is: (1) number of rooms of the hut/house, (2) a dummy for non-flammable walls, (3) a dummy for a non-flammable roof, (4) a dummy for having floors that are not bare, (5) a dummy for a toilet other than the bush, (6) a dummy for having access to electricity, (7) number of storeys, (8) land owned with legal document (lal purja), (9) ox-pulled ploughs, (10) hoes, (11) grinding machines, (12) threshers, (13) tractors, (14), sickles/scythes, (15) carts, (16) stoves, (17) pressure cookers, (18) beds, (19) lamps, (20) generators, (21) bicycles, (22) motorbikes, (23) chairs, (24) cooking pots, (25) sewing machines, (26) mobile phones, (27) radios, (28) music players, (29) television sets, (30) solar panels, (31) computers, (32) cows, (33) buffalos, (34) oxen, (35) goats.

18 This follows the guidance in Filmer and Pritchett (2001). The first principal component captures sufficient variation in the data.

19 To do this, we subtract the mean of the wealth index, and then divide by its standard deviation.

20 These results present something similar to a Difference-in-Differences specification. However, the baseline data is recalled rather than measured at baseline.

21 There are some specifications in the robustness checks where the differences between the intervention and comparison groups are statistically significant, but we do not emphasise these results as they are not supported by the main estimation procedure used for Section 6.

22 In the robustness checks, the multivariate regression specifications do, in fact, suggest that the project had a positive impact on household wealth. However, we do not emphasise this finding, because it does not emanate from the main kernel matching procedure used to generate the results for this report.

23 In the robustness checks, the sign and magnitude of the project's effect on savings does not change substantially, but the results are no longer statistically significant for the multivariate regression specifications.

24 We elected not to cluster our standard errors at the community level because this would result in a small number of clusters and would be likely to bias our standard errors downwards.

25 It should be noted that, for all these regression techniques, we report robust standard errors. However, the standard errors are not bootstrapped as in the main results in Section 6.

26 We are able to test whether the covariates are distributed sufficiently similarly for the intervention and comparison group using Rubin's (2001) tests. For the matching variables used in this report, with the kernel matching algorithm, Rubin's B = 24.0, and Rubin's R = 1.41. According to Rubin's recommendations, this suggests that the covariates are sufficiently balanced for OLS regression methods to be valid.

27 Choosing whether to match with and without 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, especially when the distribution of the propensity score is very different in the intervention and comparison group. However, allowing for replacement increases the variance of the estimates because, in effect, the number of distinct comparison observations is reduced (Caliendo and Kopeinig, 2008).

28 Following the guidance of Abadie and Imbens (2008), we calculate robust standard errors analytically using the effects module in Stata. These standard errors are not bootstrapped.

29 We estimate Equation 2 without restricting the data to the area of common support. However, we also test whether our results are sensitive to re-estimating the regressions with propensity score weighting. This makes little difference to the results shown in Appendix 5.

Oxfam Effectiveness Reviews

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

© Oxfam GB November 2017

ISBN 978-1-78748-107-7

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

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

Oxfam America (www.oxfamamerica.org)
Oxfam Australia (www.oxfam.org.au)
Oxfam-in-Belgium (www.oxfamsol.be)
Oxfam Brasil (www.oxfam.org.br)
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 IBIS (Denmark) (http://oxfamibis.dk/)
Oxfam India (www.oxfamindia.org)
Oxfam Intermón (Spain) (www.oxfamintermon.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 (Netherlands) (www.oxfamnovib.nl)
Oxfam Québec (www.oxfam.qc.ca)
Oxfam South Africa (http://www.oxfam.org.za/)