Impact Evaluation Report

Nepal – Pavitra Vegetable Seed Cooperative Project

Enterprise Development Programme (EDP)
Oxfam GB

Bet Caeyers

Global Impact Evaluation Advisor, Oxfam GB

May 2014

OXFAM

www.oxfam.org

Table of Contents

Table of figures

Table of tables

Acknowledgments

I would like to thank the staff of the Pavitra Vegetable Seed Cooperative and the Oxfam team in Nepal for being incredibly supportive during this review. Particular thanks are due to Anthony Scott Faiia, Karuna Amatya, Rigendra Khadka, Buddhi Raj Tamang, Shyam Paudyal, Mahendra Narayan Mahato, Sudha Khadka, Susan Parajuli, Suravi Gurung, Krishna Prasad Dhakal and Naul Singh Khatri.

I also greatly recognise the invaluable cooperation and support provided by the EDP team in Oxford – Hugo Sintes, Alan Doran and Maria Michalopoulou. Finally, I would like to express my gratitude to Bijendra Basnyat, the consultant who managed the baseline and endline data collection work, and Jonathan Kaminski for his assistance during data analysis.

Executive Summary

Oxfam GB's Enterprise Development Programme (EDP) is specially designed to improve the livelihoods of smallholder producers, while simultaneously promoting positive gender outcomes. The main strategy involves supporting the development of viable commercial enterprises, typically owned and run by the smallholder farmers themselves. These enterprises focus on exploiting both market opportunities based on the comparative advantages of their members, and identified potential for enabling meaningful female participation.

In Surkhet, Nepal, vegetable seed production was identified as the agricultural activity with the highest income potential and the highest potential to benefit women and smallholder farmers. In 2010, Pavitra Janakalyan Agricultural Cooperative, henceforth Pavitra, was selected by EDP as the cooperative most active in this sector. It was Pavitra's aspiration to tap into the large demand for vegetable seeds within the district, regionally, nationally and internationally. However, Pavitra on its own was not able to help its members to increase the quality and quantity of their production and increase their market share. It was lacking adequate technical resources in terms of overall management, production and marketing as well as sufficient working capital to invest strategically in the business.

Starting in October 2010, EDP provided Pavitra with support in the development and execution of a focused marketing strategy for vegetable seeds. The support included: mentoring in efficiency; improving business and production procedures (passed on to farmers through the recruitment and training of field technicians); provision of working capital to invest in the expansion of seed production and marketing services and to support the payment of salaries; advisory and brokering support to make Pavitra's production scheme more market demand-driven; the organisation of Participatory Learning Classes (PLCs) targeted at women, to raise women's awareness of their rights at various levels (eg. household, cooperative, community); and support in social mobilisation, advocacy, networking and coordination with various district stakeholders. The latter support included brokering a relationship with the Kumari bank with the aim of allowing Pavitra to provide its members with small production loans for agricultural purposes at favourable rates.

The main objective of the project was to improve the food security and the livelihoods of its members and to strengthen the empowerment of women in particular, both in the enterprise and in their respective communities and households. Evidencing the extent to which these higher-level outcomes have been achieved, as well as the various change processes at work, is a priority for Oxfam, and for EDP in particular. Creating a robust evidence base to demonstrate this link is rare in social investment models. Pavitra is the first EDP project where an attempt has been made to collect common baseline and follow-up data on a random sample of members of the intervention cooperative and on a suitable comparison population. The comparison group consists of a random sample of members of agricultural cooperatives operating in neighbouring Village Development Committees (the lowest administrative units of Nepal's local development ministry), that were active in sectors that are reasonably similar to vegetable seeds in terms of EDP's targeting criteria, i.e. activities that benefit women and smallholders and that have high potential for generating income. The sample consists of 317 households in total, comprised of 132 Pavitra members, and 185 members of comparison cooperatives.

A statistical method, combining propensity-score matching (PSM) with 'double difference', was used to estimate EDP's impact on the livelihoods of the project's direct beneficiaries. Whereas PSM allows us to control for observable baseline differences between intervention and comparison households, the double difference approach further controls for (time-invariant) unobserved differences between the two groups. In many ways, the survey instruments and sampling strategy are still in a pilot phase, and the lessons learned are being used to improve the evaluation process for all future EDP projects. Notwithstanding the caveats, the evidence available on both intermediate and final outcomes yields a set of interesting results.

The most promising findings of this study relate to EDP's impact on a range of indicators of women's empowerment. Both women and men in Pavitra households expressed significantly more positive opinions on women's economic roles in the home and in livelihoods activities. Significantly more women in the intervention group agreed with positive statements about their ability to influence affairs in their community and in the cooperative than in the comparison group. Female members of Pavitra households increased their level of self-efficacy (that is, their perceptions of their ability to influence issues and events that affect their lives) to a greater extent than the women in the comparison group. We did not observe any significant changes over time with respect to women's power to influence decisions within their household.

Although we observed improved livelihoods for the beneficiary households, when assessed through an evaluation of household consumption levels, we measured similar trends for the matched comparison group and therefore conclude that there is not yet evidence of impact of the EDP project on material well-being. We found no significant changes in the proportion of households falling below the various poverty lines, the level of food security measured by the number of meals consumed (differentiated by gender) and the variety of food consumed. We should note, however, that since the households in this area were already quite food secure at baseline and only few of them fell below the standard poverty lines, there was little scope for improvement in this respect. There was also no evidence of impact when looking at improvements in indicators of material wealth such as quality of housing or ownership of assets. In contrast, using a more subjective measure of well-being, Pavitra member households were significantly less likely to describe themselves as struggling to meet their basic needs, and more likely to describe themselves as doing well. Although subjective evidence of this sort needs to be treated with caution, it does suggest that EDP has allowed Pavitra members to gain a more positive outlook on life.

A few important qualifications need to be made when interpreting the results on 'final outcomes'. First, impacts have been examined only two years after the start of the EDP intervention, and should thus be treated as interim rather than final. Next, given the huge expansion of Pavitra membership and the number of non-member seed suppliers in the two and a half years since EDP started, the study sample drawn at baseline does not represent the overall impact of the project. Studying endline characteristics of a sample of 100 new members, we observed some significant differences with the study sample of older members (that is, those who joined Pavitra before the EDP intervention). In particular, we measured a significantly higher average value for total consumption expenditures and food consumption expenditures, and even higher scores on most of the women's empowerment indicators. In combination with a few other caveats highlighted in this report, these facts imply that the improvement in outcomes such as material well-being observed on average in this study may be an understatement of the true benefit gained from EDP by the bulk of farmers.

In order to understand how these changes or absence of changes in final outcomes came about, this study also evaluated Pavitra's performance against a set of intermediate outcomes. Specifically, we looked at indicators related to access to quality services, the adoption of improved agricultural practices, use of land and labour in agriculture, crop yields and households' revenues from the cooperative (through dividend pay-outs or sales).

Pavitra members were more likely than comparison cooperative members to report that their cooperative provided fertilisers, production loans and women's participatory classes, and that the quality of advice and technical support and the provision of agricultural inputs (such as foundation seeds) had improved. However, they were no more likely to report having *used* these services. In particular, access to production loans remains restricted to relatively few members, with only 24 per cent of the households in the study sample reporting that they took out a loan over the 12 months prior to the survey. We also noted that new members were 10 percentage points more likely than old members to have received a loan. However, the fact that 66% of households did not access any loans suggests that some barriers to loan uptake remain (on either the demand or supply side). We did not find evidence of any impact on the savings behaviour of either male or female members of the households in the sample.

To some extent, the improved quality of advice and support in agricultural practices seems to have paid off. Pavitra members are now more likely to undertake their own compost-making activities and to irrigate their fields. However, compared with the comparison group, we are not seeing differential changes in the use of animal urine and improved animal manure. The use of inorganic (chemical) fertilisers has increased more rapidly for Pavitra members than for their comparison group counterparts. We have not seen any significant differences in the type of crops or seeds cultivated by the farmers. That is, both comparison and intervention groups reported to be equally commercial in the type of products that they are cultivating. However, looking again at the endline data on our sample of 100 new members, we find that new members are significantly more likely than old members to have produced commercial crops or seeds and less likely to have used inorganic fertilisers on their farms. This evidence suggests that older members, perhaps due to their age and extensive experience, may be more reluctant to change their practices, and hence might be less likely to benefit from the advice and support provided by EDP. This could explain the absence of measured impact on the adoption of improved agricultural practices and on overall household income.

We found no significant changes over time in terms of the area of land owned, the area of land used for farming, or the area of land allocated to the production of the cooperative product. Whereas there were no changes in the number of household members working on the land, the quantity of *male* labour hired from *outside* the household did increase. However, a similar upward trend was observed in the comparison group and, therefore, we cannot attribute this change to EDP's intervention. Although we have not observed changes in terms of the area of land cultivated, a higher proportion of Pavitra members reported having experienced increased yields on their farms over the last 12 years. This effect is not statistically significant, however, and should therefore be interpreted with caution.

Surprisingly, we have not observed EDP having an impact on farmer group membership. Most striking is the finding that only four out of 100 households who have joined Pavitra since 2011 reported being a member of a farmer group at the time of the endline survey. Given that EDP explicitly encouraged Pavitra to include farmer group membership in their cooperative membership eligibility criteria; this is a disconcerting result which requires further investigation.

Pavitra members have been significantly less likely than comparison groups to receive dividends over the last year. This result could be explained by the fact that Pavitra is in the midst of investing in the construction of a processing plant and therefore its profits are currently relatively low.

Overall, we have not found evidence that EDP enabled the beneficiary group to increase the quantity or the value of sales to the cooperative to a level that would have been impossible without EDP support. These findings are supported by more subjective results obtained by asking the respondents about the extent to which the value of their sales had increased since EDP's intervention. The proportions of households that reported making increased sales were approximately the same in the intervention and comparison groups.

Based on the evaluation results, we make the following set of key programme learning recommendations:

- ***Continue to build on the significant successes in women's empowerment***
  The results clearly show that the EDP intervention has successfully supported Pavitra in the improvement of women's self-efficacy and the strengthening of women's say in the running of the affairs of their enterprises and of the community more broadly. The next step will be to find creative ways to strengthen women's influence at the household level, which – as the evaluation shows – remains limited to date.

- ***Monitor impact on livelihoods***
  The evaluation has not yet found evidence of an impact on the livelihoods of the farmers who were members of Pavitra at baseline. Although we observed significant improvements in the value of their sales and their consumption expenditures, this positive trend is similarly present among members of the comparison cooperatives in the area. If EDP's aim is to lift Pavitra members' livelihoods to a higher level – that is, a level that would not have been possible without EDP's support – it may be worth investing resources into more innovative ways to boost incomes.

- ***Investigate the reason for Pavitra members' low participation in farmer groups***
  We did not observe an impact of EDP support on farmer group membership. This result could partially be owing to the fact that there was a delay of six months before collecting baseline data. However, only 58 per cent of Pavitra's older members and only four per cent of Pavitra's newer members reported being members of a farmer group by the time of the endline survey. Given that EDP explicitly encouraged Pavitra to only consider farmers who were members of a producer group as eligible cooperative members, this result is rather disconcerting and needs further inquiry.

- ***Monitor dividend pay-outs to ensure all shareholders benefit from Pavitra's increase in profits***
  The evaluation reveals that almost none of the Pavitra shareholders received any dividends over the last year. In fact, a significantly smaller proportion received dividends compared to the period before EDP's intervention. This finding can be convincingly explained by the fact that profits have been invested in the construction of a processing plant. However, it is important to monitor the paying out of dividends in the coming years to make sure all shareholders benefit from Pavitra's increase in profits.

- ***Consider further research to investigate why farmers' access to loans remains limited***
  It is a concern that this study found that loan uptake remains limited to relatively few Pavitra members i.e. 21%. It is especially disconcerting that the older Pavitra members are least likely to obtain loans. It is unclear what is driving this result. One reason might be the interest costs. At the outset, it was expected that Pavitra would be able to provide loans at favourable interest rates. The evaluation reveals, however, that the interest rate of 16% offered by Pavitra is actually higher than the average interest rate offered by comparison cooperatives (15%). Another reason might be Pavitra's recent policy change regarding loan eligibility. Qualitative evidence gathered in the context of this review hints that Pavitra has changed its policy so that only members who own a minimum amount of land *and* who are able to present a land certificate are currently eligible for loans from Pavitra. Not only does such a policy deny access to loans to landless farmers, it could also form a loan barrier to women, especially since their intra-household bargaining power remains limited. This result is worth further investigation.

- ***Continue mentoring on improved agricultural practices***
  To some extent, the improved quality of advice and support in agricultural practices has paid off as farmers have become more likely to engage in compost-making activities and irrigation of their fields. EDP may want to consider alternative ways to encourage farmers, especially the older Pavitra members, to substitute inorganic fertilisers with organic fertilisers such as animal urine and improved animal manure.

- ***Continue to strengthen and build on the evaluation process***
  EDP should be commended on the efforts it has invested in designing and carrying out a rigorous impact evaluation. In many ways, the evaluation process adhered to in this EDP project is a considerable improvement compared to other evaluations conducted in the sector. However, as always there is scope for improvement to the survey instruments and sampling strategy. Section 4.3 in this report lists a set of caveats, mainly related to data availability. We strongly recommend that EDP carefully considers these shortcomings in order to strengthen the evaluation process for future EDP projects.

1. Introduction and Purpose

The livelihoods of the majority of the world's poor people are heavily reliant on farming. As such, improving agricultural productivity and the marketing of associated products is expected to contribute significantly to poverty alleviation efforts. In addition, given the significant involvement of women in smallholder agriculture, opportunities exist to adapt interventions to maximise positive gender outcomes.[1] At the same time, most smallholder farmers in low-income countries are likely to be land constrained, poorly linked to markets, and significantly vulnerable to risk.[2] In the wake of both the current and anticipated impacts of climate change, this latter issue is of significant concern.[3]

Oxfam GB's Enterprise Development Programme (EDP) is specially designed to improve the livelihoods of smallholder producers, with an intentional bias to simultaneously promote positive gender outcomes. The main strategy involves supporting the development of viable commercial enterprises, typically owned and run by the smallholder farmers themselves. These enterprises focus on exploiting both market opportunities based on the comparative advantages of their members and identified potential for enabling meaningful female participation. While the commercial performance of the enterprises is of critical importance, this is not an end in and of itself: Oxfam GB and other EDP stakeholders desire credible feedback on the extent to which the programme's investments are also contributing towards both poverty reduction and women's empowerment. This necessitates the application of sufficiently rigorous evaluation methods.

These evaluation methods are being pursued by the EDP programme. They entail that each new EDP project undertakes a baseline and follow-up survey to capture data on both the intervention group and an appropriate comparison group, and to analyse differential changes in their outcomes over time. This – as will become clearer below – will enable EDP's social impacts to be effectively estimated.

The Pavitra Vegetable Seed Cooperative Project is among the first EDP projects where an attempt has been made to have common baseline and follow-up data on both intervention and comparison groups. In many ways, the survey instruments and sampling strategy are still in a pilot phase, and the lessons learned are being used to improve the evaluation process for all future EDP projects. Notwithstanding the caveats – which will be listed below– the evidence available yields a set of informative recommendations that we encourage the Nepal country team to take on board to improve the roll-out of the project.

---

[1] Global Donor Platform for Rural Development. *Platform Policy Brief: The Future of Smallholder Agriculture*. No.2, October 2008.
[2] Chamberlin, Jordan. *It's a Small World After All: Defining Smallholder Agriculture in Ghana*. International Food Policy Research Institute, 2008.
[3] See, for example: Morton, John F. "The Impact of Climate Change on Smallholder and Subsistence Agriculture." *Proceedings of the National Academy of Sciences of the United States of America*. Vol. 104, No.50, December, 2007.

2. The Enterprise Development Programme in Nepal

2.1. Enterprise Development Programme (EDP)

OGB's Enterprise Development Programme (EDP) aims to tackle both poverty and gender inequality through supporting the development and growth of early-stage small and medium enterprises. Enterprise development proposals are periodically developed by Oxfam GB country teams, working together with local partners. These are submitted to Oxfam GB's Investment Committee for initial commercial viability and social impact screening. The strongest proposals are then reviewed by EDP's external Enterprise Development Board for final review and approval for inclusion in the portfolio.

Figure 1 summarises the logic of EDP's Change Model as a sequence of steps in relation to the ultimate socio-economic impacts the programme is attempting to achieve, namely poverty reduction and reduced gender inequalities.

Figure 1. EDP's Change Model's logic in words

Poverty Reduction Outcomes	Gender Outcomes
1. Targeted enterprises are supported to develop stronger, more viable business plans and are provided with requisite funding, skills training, and ongoing technical and mentoring support. 2. Market penetration of enterprises increases, and they gain more market share and/or better returns per unit sold. 3. Producers, finding it potentially advantageous to sell through the enterprises, modify their production practices to cater more effectively to the demands of enterprise buyers. 4. Producers sell more and more of their produce to the enterprises. 5. Given the enterprises' greater market share and/or better returns per unit sold, they are able to offer better prices to producers and/or provide them with greater dividends. 6. The increased income earned increases overall household purchasing power. 7. The gender outcomes achieved by the programme interact with increased household wealth to reduce gender expenditure and consumption inequality.	1. EDP's gender market mapping exercises identify enterprise development opportunities that are conducive for women's meaningful engagement. 2. Effort is also undertaken to ensure that enterprise operating modalities enable women to participate meaningfully. 3. Increased numbers of women occupy important leadership positions in the enterprises and complementary community-based institutions. 4. This, coupled with women's increased access to and control over income, improves perceptions and attitudes on the economic roles of women at household and community levels. 5. The self-confidence and self-efficacy of the women participating increases. 6. Women's improved access to and control over income also enables them to own other "strategic" assets. 7. Women possess significant leverage and self-efficacy that enables them to have a greater say in household decision-making and in the running of the affairs of their enterprises and other community institutions.

2.2. EDP in Nepal – the Pavitra Vegetable Seed Cooperative project

2.2.1. Cooperative selection process

The Pavitra Vegetable Seed Cooperative project – henceforth Pavitra – has been one of EDP's projects in Nepal since October 2010. Pavitra was selected as a cooperative eligible for EDP support through a careful screening process which started off with the consultation of various stakeholders in Surkhet district in Nepal. Through this consultation process, the EDP team identified activities in the district which had both 1) high income potential and 2) the potential to benefit women and smallholder farmers. Figure 2 represents the matrix resulting from this exercise, ranking the main economic activities in Surkhet in 2011 on the basis of EDP's priority objectives.

Figure 2. Ranking economic activities based on EDP's priority objectives

Benefit to women & small-holder farmers

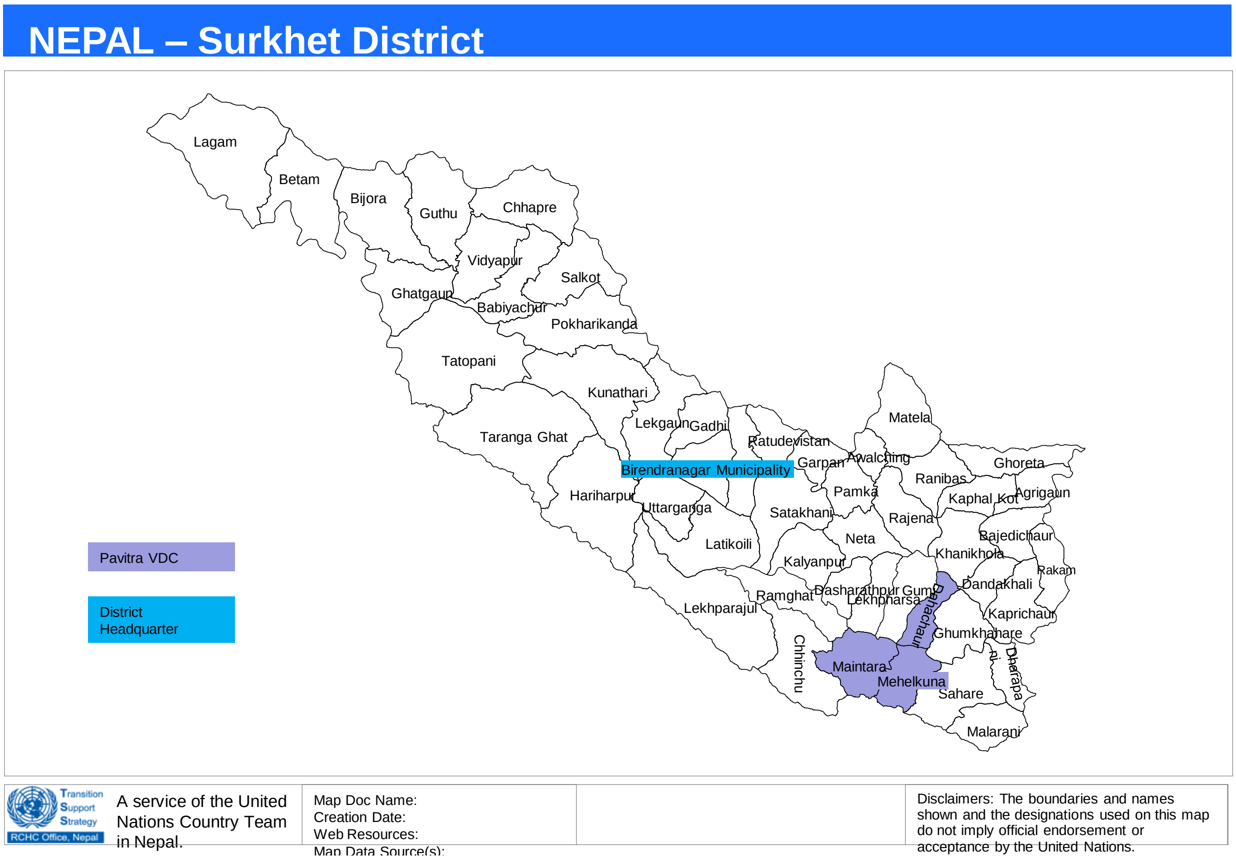
High		Dairy farming	Vegetable seeds
Medium	Honey	Fisheries	Fresh vegetables, Spices
Low	Carpentry	Goat keeping	Poultry
	Low	Medium	High

Income Potential

This process identified vegetable seed production as the economic activity in the targeted region with the highest potential to both increase incomes and to benefit women and smallholder farmers. Pavitra was eventually selected on the basis that, although operating only on a small scale, it was more active in vegetable seed production than any other cooperative in the area. The impact assessment below takes into account baseline differences between Pavitra members and comparison households as a result of this specific targeting process.

2.2.2. Cooperative background

Figure 3. Area of Pavitra's operations at baseline (2011)



Pavitra (named after a local market in the area) was established in 2001 by a small group of farmers from the Village Development Committee (VDC) of Mehelkuna in Surkhet district[4]. Prior to EDP support in 2010, Pavitra had 216 members (31% of whom were women). Pavitra was operational in three VDCs in the district, namely Dachachaur, Mehelkuna and Maintara. Figure 3 indicates Pavitra's area of operations prior to the EDP intervention.

From around 2005, some progressive Pavitra members, encouraged by support from the national NGO CEAPRED, gradually moved to small scale production of vegetable seeds. The bulk of Pavitra members, however, kept their production exclusively focused on fresh vegetables, mainly for the purpose of subsistence. At the outset of the project, the cooperative members – of whom only a small number were organised into farmer groups – were marketing an annual volume of approximately 50 tonnes of vegetable seeds through the cooperative. The average annual income of members from seed production was NRs. 15,000 (GBP 128). Seeds were marketed in 30kg sacks to a limited number of seed buyers based in Kathmandu who graded and repackaged the seeds, selling them under their own brands to the national and international markets.

[4] The Village Development Committee (VDC) is the lowest administrative unit of Nepal's local development ministry. Each district in Nepal has several VDCs, which in turn is sub-divided into wards.

Until 2001, Pavitra mainly played the role of mediator between the producers and the market. National NGOs such as CEAPRED distributed foundation seeds and fertilisers directly to individual farmers and Pavitra sold their produce to local seed companies. Market demand of vegetable seeds was largely ignored throughout the production process. Farmers made independent decisions about the type and quantity of their produce (mainly fresh vegetables), selling this on to Pavitra after harvesting, irrespective of whether their supply would meet market demand.

Farmers' technical knowledge of improved agricultural practices, such as the use of organic composting methods and information on the variety of commercial products, was limited. Pavitra provided some occasional technical advice to their individual members through a Junior Technical Assistant (JTA), but this person was neither properly trained nor paid for his/her services.

Moreover, in 2010 the cooperative was poorly managed. For instance, there was no systematic recording of shareholder membership or sales payments made to the farmers. The cooperative did not have a formal system in place to provide loans and savings services to their members. Although an occasional loan was provided to some farmer members, the number of credit beneficiaries was very small and no clear regulations were followed.

It was Pavitra's aspiration to tap into the large demand for vegetable seeds within the district, regionally, nationally and internationally. However, Pavitra on its own was not able to help its members to increase the quality and quantity of their production and increase their market share. It was lacking adequate technical resources in terms of overall management, production and marketing as well as sufficient working capital to invest strategically in the business.

2.2.3. EDP intervention

EDP provided its first grant to Pavitra in October 2010. The initial EDP investment was focused on supporting a rigorous governance structure. Its mandate was to simultaneously provide technical assistance in marketing and production and facilitate access to 'business loans' in partnership with an established financial institution, the Kumari Bank. At the outset, it was envisaged that grant support would focus on technical assistance in key areas of market development: how to add value, branding, management, and market chain improvement(including production, sorting, cleaning, grading, packaging etc). It should be noted, however, that at the time of the impact evaluation in 2013, Pavitra had not yet engaged in value- adding activities such as branding and packaging. The evaluation therefore focuses on the impact of the various EDP activities that had actually been implemented at the time of the endline data collection visit.

One important form of OGB's support to Pavitra over the last three years has been continuous **mentoring.** Areas of advisory support include the development of a clear cooperative policy (e.g. loan policy, employment policy), the systematisation of documentation procedures (e.g. computerised tracking of sales and credit payments), the design of feasible business plans, and technical advice and training on profitable seed production methods (passed on to farmers through field technicians – see below).

Furthermore, EDP provided Pavitra with working capital to invest in the expansion of seed production and marketing services (for instance, through the purchase of office equipment, irrigation infrastructure and production machinery such as grinding machines and biogas plants) and to support the payment of salaries (tapered over three years).[5] Among other things, this support allowed Pavitra to recruit three full-time Junior Technical Assistants (JTAs) whose objectives were to regularly visit the farmers in the intervention VDCs to prepare production contracts, to facilitate farmer group trainings (e.g. on organic farming), to check the quality of the seeds, to provide technical advice throughout the cultivation process (e.g. on the use of pesticides), and to settle loans and process new loan applications. JTAs were also in charge of organising farmers into farmer groups – because cooperative membership has become conditional on farmer group membership since EDP's intervention – and informing them about the benefits of a Pavitra membership.

EDP also provided advisory and brokering support to make Pavitra's production scheme more demand-driven. Whereas prior to EDP support, farmers produced goods independently from market conditions, the trading relationship between Pavitra and the farmers is now intended to be entirely demand-driven. The way this works in practice is that Pavitra board members are advised to collect information about the demand from the seed market and fix prices with regional seed companies ahead of the farming season. Equipped with this information, JTAs go to the field and arrange production contracts for the next season, stipulating the area of land that will be cultivated by each farmer, the type and quantity of seed varieties that will be produced, the price at which it will be sold, etc. Based on these data, Pavitra calculates the amount of foundation seeds that are needed to produce the specified amounts which, as per contract, get sold to Pavitra after harvesting.

Complementing the direct support to business activities, EDP, through its partner organisation Environment Development Society (EDS), has organised **Participatory Learning Classes (PLCs)** targeted at female farmers living in the intervention area. These classes, run several times a week in various neighbourhoods, are facilitated by women recruited from the intervention communities and trained by EDS. Topics covered include issues related to women's rights, cooperative membership rights (e.g. access to dividends), guidelines on how to report complaints, and community issues (e.g. WASH). The objective of these PLCs is to raise awareness among women of their rights – at various levels (household, cooperative, community) – and to inform them of how they can protect these rights.

Finally, Oxfam and EDS provided **support in social mobilisation, advocacy, networking and coordination** with district stakeholders such as the District Agricultural Development Office, District Cooperative Office, District Development Committee, seed trading companies and agro vets. Oxfam also brokered the relationship between Pavitra and the Kumari Bank, which resulted in the provision of commercial loans to Pavitra, who in turn are now in a position to provide small production loans for agricultural purposes (currently limited to NPR 10,000 – NPR 90,000 per farmer) to their members at favourable rates.[6] Note that loans are agreed at the level of the farmer group and a new loan can only be requested once all members of the farmer group have settled their loans with Pavitra. Moreover, Pavitra has recently changed its credit policy so that now only members who own a minimum amount of land and who can present a land certificate can get access to loans.

[5] Note that prior to EDP support most cooperative staff were working on a voluntary basis.

[6] On paper, Pavitra borrows money from Kumari bank against 12 per cent interest rate and lends to farmers at 16 per cent interest rate.

2.2.4. Expected project beneficiaries

We expect there to be three types of potential beneficiaries of this particular EDP intervention. These are:

1) **Pavitra shareholder members**. Farmers who are members of Pavitra benefit, in theory, from the full range of services outlined above, including trading services, access to savings facilities, loans, subsidised fertilisers, foundation seeds on credit and technical advice. In addition, Pavitra members are eligible to dividends whenever these are paid out. We further distinguish between 'old members' who were members of Pavitra prior to the EDP intervention and 'new members', who only became members after EDP had started providing support.
2) **Non-member seed suppliers**. Farmers who are not members of Pavitra but who sell their produce to Pavitra benefit from a similar set of services as Pavitra members (including the provision of fertilisers and foundation seeds, and technical advice), except that they do not have access to formal training, loans and savings services from Pavitra.
3) **Indirect beneficiaries**. It is likely that other community members – such as labourers hired to work on the farms, neighbouring farmers and/or members of the same farmer group, indirectly benefit from the EDP intervention, via the transfer of information on agricultural practices through word of mouth. Also, farmers who buy the vegetable seeds benefit from the good quality since it improves the quality and quantity of fresh vegetables they produce.

Unfortunately, as will be discussed below, the data available only allow us to rigorously assess the impact on the first category of community members, i.e. the Pavitra shareholder members. More specifically, we are only able to look at the impact on 'old members', i.e. those farmers who were already a member of Pavitra in 2010. We are not in a position to look at the impact on either new members or non-member seed suppliers, or on indirect beneficiaries in the communities. This important caveat has been taken into account when interpreting the impact assessment results below.

2.2.5. Expected intermediate and final project outcomes

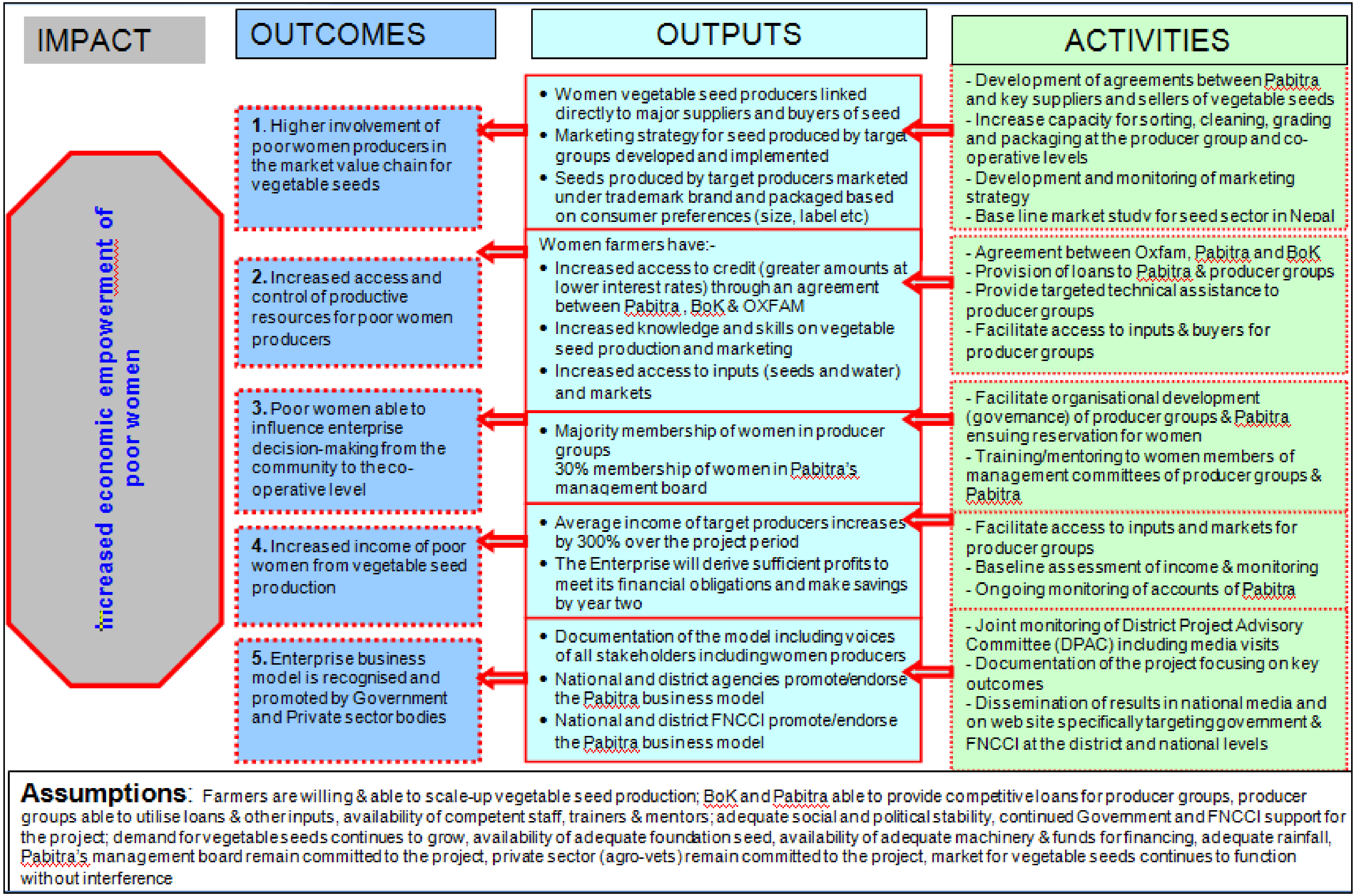
Figure 1 above outlined the general logic model through which EDP was expected to deliver its desired results, i.e. to improve the food security and livelihoods of the enterprise's members and to strengthen women's empowerment in particular, both in the enterprise and in their respective communities and households. Figure 4 illustrates EDP's change model specific to Pavitra.

This report assesses the impact of the project on different stages of the results chain, i.e. the intermediate and final outcomes. Table 1 provides an overview of the particular set of outcomes that the project has been evaluated against in this review. Section 5 below contains more details on how these outcomes were measured and tracked over time.

Table 1. List of expected intermediate and final project outcomes

Outcome type	Outcome area	Outcome
Intermediate outcomes	Access to quality services	Access to training in improved agricultural methods Access to advice and technical support Access to agricultural inputs (water, foundation seeds, etc.) Access to fertilisers Access to credit and savings Access to women's participatory learning classes
	Adoption of improved agricultural practices	Production of commercial crops/seeds Use of inorganic fertilisers Engagement in compost-making activities Use of animal urine for compost-making Use of improved animal manure Use of irrigation techniques
	Agricultural use of land and labour	Land ownership Area of land used for farming Proportion of land allocated to farming Number of male and female household members involved in agriculture Hired employment on farms Membership in farmer groups
	Yields	Yields on farms
	Revenue from cooperative	Revenues from dividend pay-outs Revenues from sales of produce to cooperative and other merchants
Final outcomes	Improved livelihoods	Food- and non-food consumption expenditures Poverty rates Ownership of assets
	Food security	Food security score Diversification of food consumption
	Women's empowerment	Men's and women's perceptions of women's economic roles Intra-household decision-making power of women Influence of women in the community Influence of women in the cooperative Self-efficacy of women

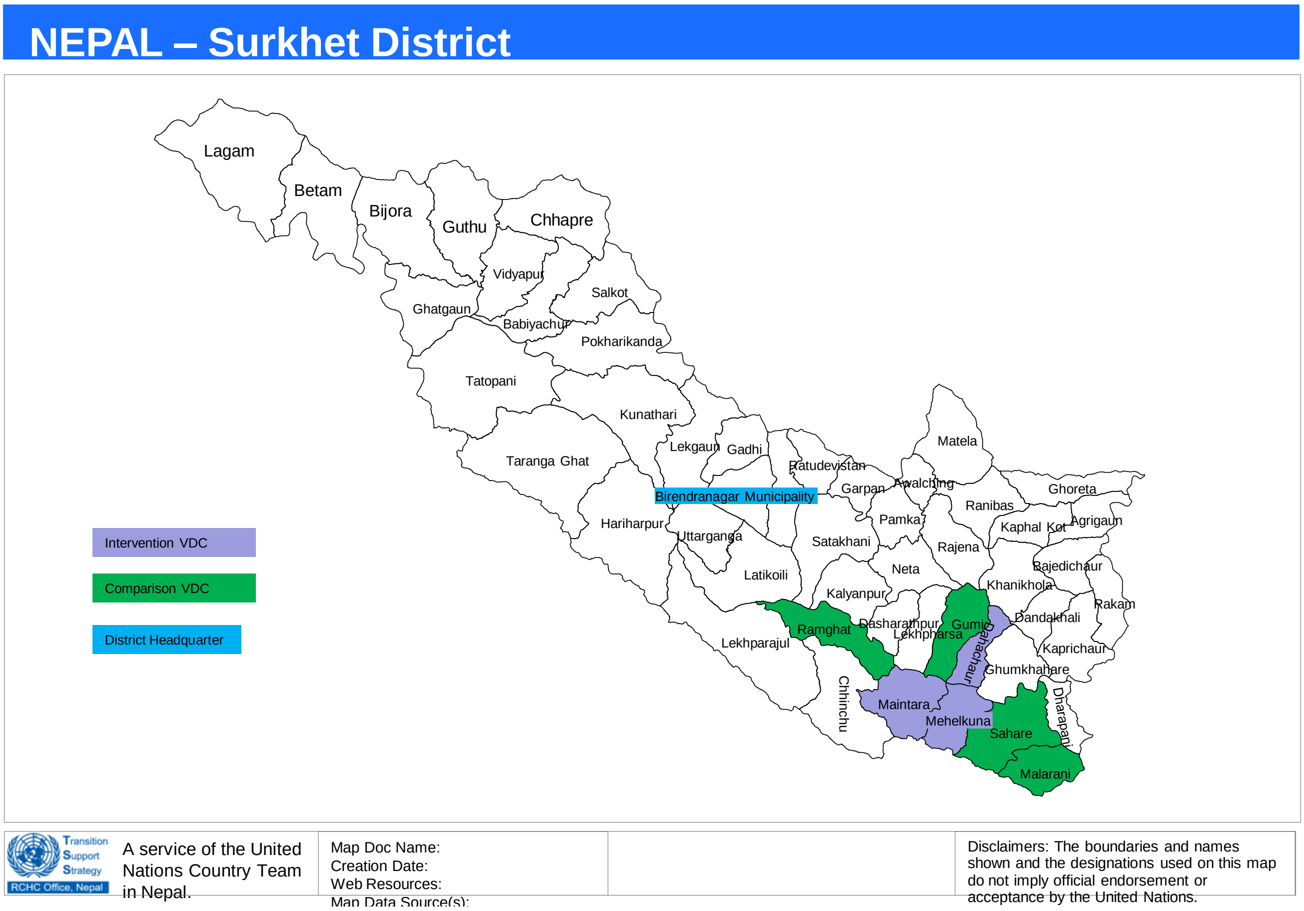
Figure 4. EDP Pavitra Nepal Change Model



3. Evaluation Approach

A much more detailed explanation of the choice of evaluation approach in this study can be found in Appendix A1. In summary, the core challenge of any impact evaluation is to credibly estimate the net effect of an intervention or programme on its participants. The randomised experiment is regarded by many as the most rigorous way of estimating project impact. The EDP Pavitra project in Nepal however forms a clear example of a setting where it was not possible to randomly select beneficiaries. As discussed in Section 2.2.1, the producer group being targeted was purposefully selected, with Pavitra being identified as an eligible cooperative through a careful consultation process, and individual participants, in turn, being self-selected following Pavitra's membership eligibility criteria. As such, participation was not random and we had to identify appropriate non-random comparison groups for the purpose of this study. Specifically, we identified three agricultural cooperatives in neighbouring VDCs that were active in sectors that are most similar to vegetable seeds in terms of EDP's top-level targeting criteria, i.e. activities that 1) benefit women and smallholders and 2) have high income potential. Appendix A1 provides more details about the selection and validity of the comparison group used in this study. Combined with the beneficiary households, we obtained a study population with households from three intervention VDCs and four comparison VDCs, the locations of which are indicated on Figure 5.

Figure 5. Locations of intervention and comparison VDCs



Despite the comparison group being carefully selected to be as similar as possible to the intervention group (given that the comparison group was non-random, and taking into account known differences such as the sector), the two groups may have differed in a) their pre-programme outcomes; and b) their other observed and unobserved characteristics. As such, directly comparing the endline outcome differences between the intervention and comparison groups could well have lead to biased intervention estimates. To address these issues, we used a statistical method which combines propensity score matching (PSM) techniques with double difference. Appendix A1 explains this procedure in more detail. Whereas PSM allowed us to control for observable baseline differences between intervention and comparison households, the double difference approach further controlled for (time-invariant) unobserved differences between the two groups. In short, this was achieved by first matching intervention households with those comparison households that were observationally similar to them at baseline (using a particular matching algorithm) and then, in a next step, by comparing the changes in the outcomes of the intervention households with the equivalent changes in outcomes of their respective matches. In addition, we used Ordinary Least Squares (OLS) regression techniques to test the robustness of our results to the controlling for changes in observable characteristics over time.

4. Data

4.1. Data source and sample

Baseline data were collected in April 2011, i.e. 6 months after EDP support was launched, but very close to the start of the new activities laid on by Pavitra.[7] The baseline sample consisted of 136 intervention households and 210 comparison households, each set randomly selected from the membership list compiled through consultation with the cooperative representatives. Note that, as required for PSM, we sampled a significantly greater number of comparison households than intervention households, to enhance the probability of finding good matches.

In August 2013 (at endline), we revisited 132 of the 136 intervention households and 185 of the 210 comparison households, obtaining a study sample of 317 respondents in total. Table 2 provides a summary of this sample:

Table 2. Summary of the study sample

Intervention group (Pavitra members)		132
Comparison group (members of comparison cooperatives)	Bheri Krishi	102
	Bheri Pashu Tatha Krishi	47
	Bhaktapur	36
	Total number of comparison households	185
TOTAL sample		317

[7] During the first six months, the main activity that had been carried out was farmer group formation. The contract with Kumari bank was only signed on 10th February 2011. No loans had been given out to members by the time of the baseline study. Pavitra had provided foundation seeds, but harvesting was not done until after baseline.

In each data collection round, three types of questionnaires were administered to all sampled households: a general household questionnaire for cooperative members; a specific questionnaire for the male household member who was most informed about agricultural practices; and a specific questionnaire for the female household member who was most informed about agricultural practices. Each time, the gender of the enumerator was matched to the gender of the respondent.

At endline, we asked the respondents exactly the same questions that were asked at baseline. This allowed us to study changes over time for both intervention and comparison groups. In addition, we asked the respondents a set of 'recall' questions on topics that had not been covered at baseline, such as yields and access to credit. Unfortunately, the lack of baseline data prevented us from using the more rigorous difference-in-difference approach for the evaluation of project impact on these indicators. Nevertheless, a comparison of endline answers between intervention households and their respective matches in the comparison group yielded some informative results.

In addition, at endline we also visited the households of 100 'new members' (i.e. individuals who joined Pavitra after baseline) and 100 'non-member seed suppliers' (i.e. individuals who sell their seeds to Pavitra but who are not shareholder members). Since no baseline data was available for these households, we could not use the difference-in-difference approach to rigorously assess the impact the project has had on these beneficiary groups. Restricted to endline data, and with no control groups to compare these outcomes to, Section 6 below provides some descriptive statistics on these members to give a tentative idea of the potential change in the profile of Pavitra members since baseline and of the potential impact EDP has had on these other beneficiaries.

4.2. List of caveats

Before we present the results, it is important to recognize a set of caveats, mainly related to data availability. These constraints have to be borne in mind when interpreting the findings below:

1. Impacts measured only two years after the start of the EDP intervention have been examined, and should thus be read as interim rather than final.
2. It is crucial to note the unavoidably unrepresentative nature of the sample at endline, given the huge expansion of Pavitra membership (from 216 members at baseline to 816 members at endline) and the increased number of non-member seed suppliers (from 216 to 1340) in the two and a half years since EDP started. Therefore, the improvement in outcomes such as material well-being observed on average for this study is quite likely to be an understatement of the true benefit derived from EDP by the bulk of farmers.
3. The baseline data collection was delayed until six months after the start of the project and therefore we might be underestimating the overall impact on some of the outcome indicators for which changes came about at early stages of the project (see Section 4.1).
4. Whereas the data available has allowed us to evaluate the *impact* of the project on the beneficiaries targeted by the project, it does not allow us to look at the *targeting effectiveness* of the project. That is, we are not in a position to answer questions such as 'are we reaching the right people?'
5. The data did not allow us to capture the impact on indirect beneficiaries, such as non-member farmers selling their produce to Pavitra and/or other community members (see Section 2.2.4).

6. No baseline data are available on various important intermediate outcome indicators, such as sales, credit and yields. This means that we have not been able to use the more rigorous difference-in-difference impact evaluation tools to estimate the impact on these indicators.
7. It is not possible to entirely exclude the possibility that the impact of EDP has spilled over to comparison households, for instance through the transfer of information on improved agricultural practices between cooperative leaderships. Although this indirect effect is likely to be substantially smaller than the direct effect on Pavitra members, this caveat does imply that we may be underestimating the true impact of EDP.
8. Given the relatively small sample size (132 intervention households), we may not be able to demonstrate the statistical significance of small changes. Therefore, 'no evidence of impact' does not necessarily mean that there were no minor changes in reality. Moreover, owing to the small sample it has not been possible for us to conduct meaningful heterogeneity analysis, i.e. to look at whether some sub-groups benefitted more or less so than others. This means that, where we find no evidence of impact *on average*, it is not possible to investigate whether this is true for all sub-groups.
9. Some indicators chosen by EDP to track improvement over time are those on which the beneficiary group scored already very well at baseline. For instance, as documented below, most Pavitra members were already food secure and very few fell below the standard poverty lines. This implies that there was little scope for improvement. Therefore, evidence of impact on such indicators should not necessarily be interpreted as a negative result. However, it does raise the question of whether appropriate indicators were selected at baseline and/or whether the right people were targeted.

5. Main results

5.1. Guide to the interpretation of results

We have attempted to write a report that is free of excessive technical jargon. However, there are a few statistical concepts that it is useful to introduce before reviewing the results. Appendix A2 provides a detailed guide to the interpretation of the results, including the definition of 'significant impact' and a detailed explanation of how to read the tables and how to derive conclusions from them. Readers who are less familiar with quantitative data analysis are strongly encouraged to read these guidelines before moving on to the results sections below.

5.2. Results on intermediate project outcomes

5.2.1. Access to quality services

No baseline data are available on households' former access to services such as technical advice or credit. For these intermediate outcomes we are therefore restricted to endline data for a comparison of matched intervention and comparison households. If there were no significant differences between the two matched groups at baseline, this constraint would not be problematic. To verify this, we asked the respondents the recall question of whether or not the service was provided in their community prior to 2011. The results presented in Panel A in Table 3 confirm that there were no significant differences in the responses provided by intervention and comparison households. Hence, acknowledging that this ignores potential differences in terms of the quality of the services at baseline, we consider a comparison of endline data on the matched sample to be informative about the potential impact of the EDP intervention.

Table 3. Access to quality services

			Training in improved agricultural methods (1)	Advice and technical support (2)	Provision of agricultural inputs (3)	Provision of fertilisers (4)	Provision of production loans (5)	Women's participatory learning classes (6)
Panel A	% households reporting that service was being provided prior to 2011	Intervention	58.3	67.4	65.9	56.8	53.0	42.4
		Comparison	60.3	59.5	61.4	62.4	47.2	31.6
		Difference	-1.9	7.9	4.5	-5.6	5.8	10.8
			(8.2)	(7.6)	(8.2)	(7.6)	(8.4)	(7.8)
Panel B	% households reporting at endline that cooperative provided service over the last 12 months	Intervention	78.8	87.1	87.1	87.9	76.5	53.0
		Comparison	74.6	81.5	78.7	75.3	59.9	35.0
		Difference	4.2	5.7	8.4	12.6*	16.6**	18.1**
			(6.5)	(6.1)	(6.1)	(6.5)	(7.8)	(7.6)
Panel C	% households reporting that service provided was very good (rather than 'good' or 'not good')	Intervention	27.2	40.6	38.3	32.7	23.6	20.4
		Comparison	20.2	23.6	23.8	22.5	23.0	9.3
		Difference	7.0	17.0**	14.5*	10.2	0.6	11.1
			(8.0)	(7.7)	(7.5)	(8.0)	(10.3)	(11.3)
Panel D	% households reporting to have made use of service provided by cooperative over the last 12 months	Intervention	62.9	75.8	81.1	81.1	55.3	41.7
		Comparison	71.0	78.7	76.5	71.4	48.8	31.7
		Difference	-8.1	-2.9	4.5	9.7	6.5	9.9
			(7.2)	(6.5)	(6.5)	(6.9)	(8.1)	(7.6)
	# matched intervention observations		132	132	132	132	132	132
	# observations total sample		299	299	299	299	299	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

We asked both Pavitra members and members of the comparison cooperatives whether a range of serviced were provided by their cooperative over the previous 12 months and whether or not anyone in their household had made use of the service over the same period. For each service, we also asked those respondents who had reported to have used the service whether they thought the quality of the service provided was 'not good', 'good' or 'very good'.

Panel B in Table 3 suggests that, relative to the members of the comparison cooperatives, Pavitra members were more likely to report that their cooperative provided fertilisers (13 percentage points more likely), production loans (17 percentage points more likely) and women's participatory classes (18 percentage points more likely). We have not found significant differences in the provision of training in improved agricultural methods, provision of advice and technical support, and delivery of agricultural inputs (such as foundation seeds). However, we have found evidence of differences in terms of the *quality* of these services provided. Whereas only 24 per cent of the comparison group reported that the quality of advice and technical support provided by their cooperative was 'very good', this number was 41 per cent for the intervention group. Similarly, 24 per cent of comparison respondents reported that the quality of the agricultural inputs provided was 'very good', compared to 38 per cent of the intervention respondents.

Table 4. Credit and savings

		% households reporting to have accessed loans (1)	Interest rate (2)	Savings – male (3)	Savings – female (4)
Baseline	Intervention			159.091	170.303
	Comparison			62.669	63.352
	Difference			96.421 (62.490)	106.951** (44.713)
Endline	Intervention	24.2	15.94	220.909	265.947
	Comparison	30.9	15.43	142.392	216.558
	Difference	-6.7 (7.0)	0.59** (0.28)	78.517 (55.957)	49.389 (51.404)
Change	Intervention			61.818	95.644
	Comparison			79.723	153.206
	Difference (PSM diff-in-diff)			-17.905 (45.302)	-57.563 (51.948)
OLS with matching weights				-14.700 (49.396)	-63.816 (53.487)
# matched intervention observations		132	32	132	132
# observations total sample		299	94	299	299

Notes: No baseline data available for columns 1-2; averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level

Despite the improved *provision* of services, beneficiary members were not more likely than control members to report having *utilised* such services as provided by their cooperative. Access to production loans, in particular, seems to be an area with scope for improvement. In a separate question we asked the respondents whether anyone in their household had borrowed any money from the cooperative over the last 12 months. The results are presented in column 1 of Table 4. Whereas 31 per cent of the comparison group had obtained loans from their cooperative over the previous 12 months, only 24 per cent of the intervention group reported to have accessed credit (note, however, that this negative difference is not statistically significant).

This somewhat negative result could be explained by the additional finding that the interest rate charged by Pavitra was slightly higher than the one charged by the comparison cooperatives, i.e. approximately 16 per cent versus 15 per cent on average – see column 2 in Table 4. Moreover, during qualitative discussions undertaken for this review, project beneficiaries highlighted concerns regarding the credit regulations being imposed by Pavitra at that time. In particular, according to various respondents, only members who owned a minimum amount of land *and* who were able to present a land certificate were able to obtain loans from Pavitra at that time (this finding was also confirmed by the Pavitra leadership). This result is worth further investigation.

Finally, we asked the household respondents, by gender, about the total amount of savings or contributions to savings associations - such as the cooperative – that they had made over the previous 12 months. We have not found any significant differences between intervention and comparison households in that respect - see columns 3-4 in Table 4.

5.2.2. Adoption of improved agricultural practices

To some extent, the improved quality of advice on, and support for, agricultural practices seems to have paid off. As shown in Table 5, Pavitra members were eight percentage points more likely to report having undertaken compost-making activities over the previous 12 months, and nine percentage points more likely to have irrigated their fields. However, we have not observed significant differences between intervention and comparison groups regarding changes in the use of animal urine and improved animal manure. Moreover, the use of unfavourable inorganic (chemical) fertilisers has increased more rapidly for Pavitra members compared to their counterparts in the comparison cooperatives. We have not observed any significant differences in terms of changes in the type of production crops or seeds cultivated by the farmers. Both comparison and intervention groups have, to an equal degree, become more commercial in the type of products that they are cultivating.

Table 5. Adoption of improved agricultural practices

		% Production commercial crops/seeds (1)	% Inorganic fertiliser consumption (2)	% Compost-making (3)	% Use of animal urine (4)	% Use of improved animal manure (5)	% Irrigation (6)
Baseline	Intervention	78.0	79.5	61.4	50.0	34.8	75.0
	Comparison	84.1	85.7	68.3	66.9	48.7	72.6
	Difference	-6.0 (5.1)	-6.2 (5.2)	-6.9 (7.3)	-16.9** (7.4)	-13.8* (7.9)	2.4 (6.0)
Endline	Intervention	86.4	89.4	69.7	53.0	40.9	89.4
	Comparison	93.2	91.3	72.0	72.1	54.7	76.2
	Difference	-6.9* (3.9)	-1.9 (4.3)	-2.3 (7.0)	-19.1*** (7.2)	-13.8* (7.4)	13.2** (5.7)
Change	Intervention	8.3	9.8	8.3	3.0	6.1	14.4
	Comparison	9.1	5.6	3.7	5.2	6.0	3.6
	Difference (PSM diff-in-diff)	-0.8 (4.7)	4.3 (3.9)	4.6 (4.6)	-2.1 (3.6)	0.0 (6.5)	10.8** (4.8)
OLS with matching weights		-2.8 (4.8)	6.0* (3.1)	8.3** (3.6)	-2.0 (3.2)	-4.6 (6.3)	8.5** (4.1)
# matched intervention observations		132	132	132	132	132	132
# observations total sample		299	299	299	299	299	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

5.2.3. Agricultural use of land and labour

Table 6 indicates no significant changes over time, for either intervention or comparison groups, in the area of land owned (column 1), the area of land used for farming (column 2) or the area of land allocated to the production of the cooperative product (column 3).. We also observed little to zero change in the number of male or female household members who spend at least half of their time doing agricultural activities. Given that this number already stood quite high at approximately two males and two females per household at baseline, there probably was little scope for an increase in this number. Instead, we would have expected there to be an increase in the quantity of labour hired from outside of the household. Indeed, Pavitra members experienced an increase in the quantity of casual or full-time male labour hired to work on the farm, i.e. an increase of approximately two men days on an annual basis (an increase from 1.7 to 4 working days per year). Given the similar trend observed for the comparison group, however, we cannot attribute this change to the EDP intervention. Whereas both the beneficiary group and the non-beneficiary group experienced a significant increase in the number of females hired since baseline, this increase was relatively more pronounced for the comparison group than for the intervention group (an increase from 2.5 to 4.3 working days for the intervention group and an increase from 0.5 to 5.4 for the comparison group). In conclusion, there is no evidence that suggests that EDP has created more jobs on the farms.

Table 6. Agricultural use of land and labour

		Area of land owned (hectares) (1)	Area of land used for farming (hectares) (2)	% farmed land allocated to cooperative product (3)	# male household members involved in agriculture (4)	# female household members involved in agriculture (5)	On-farm men-day hires per year (6)	On-farm women-day hires per year (7)	% households that mainly depend on agriculture (8)	% member of farmer group (9)
Baseline	Intervention	0.423	0.469	21.8	1.543	1.720	1.705	2.515	43.2	52.3
	Comparison	0.419	0.450	25.8	1.590	1.689	0.763	0.456	45.7	62.1
	Difference	0.004	0.020	-4.0	-0.050	0.031	0.942**	2.060***	-2.6	-9.8
		(0.025)	(0.026)	(3.0)	(0.263)	(0.241)	(0.456)	(0.437)	(5.4)	(7.8)
Endline	Intervention	0.433	0.467	22.6	1.318	1.652	3.985	4.273	76.5	57.6
	Comparison	0.410	0.451	26.9	1.464	1.461	2.911	5.417	62.4	66.7
	Difference	0.024	0.016	-4.3	-0.146	0.191	1.074	-1.144	14.1*	-9.2
		(0.029)	(0.026)	(3.2)	(0.158)	(0.133)	(1.547)	(1.311)	(7.8)	(7.9)
Change	Intervention	0.010	-0.003	0.8	-0.244	-0.068	2.280	1.758	33.3	5.3
	Comparison	-0.009	0.001	1.1	-0.169	-0.228	2.149	4.961	16.6	4.7
	Difference	0.019	-0.004	-0.3	-0.100	0.160	0.132	-3.204**	16.7*	0.6
	(PSM diff-in-diff)	(0.020)	(0.011)	(1.5)	(0.303)	(0.236)	(1.665)	(1.352)	(9.6)	(3.7)
OLS with matching weights		0.014	-0.006	-1.1	-0.205	-.0128	0.256	-3.332**	18.4*	0.4
		(0.014)	(0.008)	(1.2)	(0.241)	(0.175)	(1.468)	(1.618)	(10.5)	(3.6)
# matched intervention observations		132	132	132	132	132	132	132	132	132
# observations total sample		299	299	299	299	299	299	299	299	299

Notes: Averages are calculated after matching; Standard errors are shown in parentheses; * significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level.

Column 8 in Table 6 reveals that Pavitra members have become significantly more dependent on agriculture since baseline, with 33 percentage points more members reporting that at least half of their household's livelihood depends on crop cultivation, livestock rearing and/or the production of value-added agricultural products (an increase from 43 per cent at baseline to 76 per cent at endline). Comparing this trend to the one experienced by the comparison households, we can attribute 18 per cent of this increased dependence on agriculture to the EDP intervention.

Finally, we have not found evidence of any significant impact of EDP on membership in farmer groups (column 9). However, this result should be treated with much caution given that, as discussed in Section 4.3, the baseline data were possibly collected some time *after* the farmer group formation had started.

5.2.4. Yields

Vegetable seed production is expected to be more profitable than the production of other crops as one unit of land can produce more seeds than crops. This implies that even small land holders could potentially see their profits increase after a switch to seed production. Also, farmers who were already producing vegetables seeds prior to EDP support could see their yields increase thanks to better access to information about improved agricultural practices.

Unfortunately, the baseline survey did not collect any information on yields. At endline, we asked the respondents about the extent to which the total yield from the cultivation of the cooperative product had changed over the previous two years. Table 7 shows that a higher number of beneficiary members reported having experienced increased yields on their farms. Although the magnitude of the estimated difference seems relatively high, i.e. nine percentage points, it is not statistically significant. Although this evidence is consistent with EDP having a potential impact on yields, we have refrained from drawing any firm conclusions in this respect.

Table 7. Yields

		% households reporting no change in yields (%) (1)	% households reporting an increase in yields (%) (2)
Endline	Intervention	18.9	81.1
	Comparison	27.4	72.6
	Difference	-8.5	8.5
		(6.7)	(6.6)
# matched intervention observations		132	132
# observations total sample		299	299

Notes: No baseline data available; Averages are calculated after matching; Standard errors are shown in parentheses; * significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level

5.2.5. Revenue from cooperative (dividends and sales)

With the support of the partner organisation EDS, it was envisioned that farmers (mainly women) would adopt commercial vegetable seed production in their own producer groups, and that Pavitra would help to raise their production quantity and quality. The project was expected to increase the market share of Pavitra significantly by directly accessing a greater range of buyers (locally, nationally and, at a later phase, internationally).

Table 8 presents the results concerning the revenue generated from cooperative dividend payouts and the quantity and value of the produce sold to the cooperative. Again, no baseline data was available for outcomes captured in columns 4-8. We should therefore note that these particular results might have been driven by unobserved differences at baseline that we were not able to control for.

Column 1 reveals that, as a result of the EDP intervention, Pavitra members had become 41 percentage points *less* likely to have received dividend payouts in the year preceding the endline data collection visit. At endline, only five per cent of the Pavitra members reported to have received dividends over this 12 month period.[8] This result can be explained by the fact that Pavitra was in the midst of investing in the construction of a processing plant. Its members had been informed that no dividends would be paid out during this investment period.

Overall, we have not found clear evidence that the EDP intervention enabled the beneficiary group to increase the quantity or the value of sales to the cooperative to a level that would have been impossible without EDP support. The proportion of households selling to the cooperative and the proportion of total produce sold to the cooperative did increase by 41 percentage points and 35 percentage points respectively, but this increase was not statistically significantly different from the increase experienced by the comparison farmers once we take into account the effect of observed changes in other characteristics (through OLS with matching weights).

Similarly, we have not observed significant differences in terms of the reported gross value of total sales to the cooperative (column 5) on average for all farmers ('All') and on average for those farmers reporting to have sold their produce to the cooperative ('Sellers'). At the outset, average annual income from sales was expected to grow from NPRs 15,000 (GBP 90) to NPRs 45,000 (GBP 271). Table 8 suggests, however, that the average sales of those reporting to have sold to Pavitra ('Sellers') currently stands at NPR 14,186, which is surprisingly small given the high expectations at the launch of EDP. Interestingly, Pavitra members were significantly (i.e. 34 percentage points) *less* likely to sell produce to merchants other than Pavitra (column 6), and the gross value of these sales was significantly smaller (column 7). We have not observed any significant differences in the total sales of the cooperative product to any merchant (cooperative or otherwise). Although it is quite likely that these findings are driven by unobserved differences at baseline which we were not able to control for (see discussion above), the observation that comparison households are generating similar or even more revenues from sales at endline than Pavitra members suggests that EDP has not yet achieved the levels of sales that it was aiming for.

[8] In theory, when a dividend is declared then it is to be distributed to all shareholder members. It is therefore unlikely that some beneficiaries received dividends and others did not. It is possible that the five per cent of respondents reporting to have received dividends suffered from recall bias.

Table 8. Revenue from cooperative (dividends and sales)

		% households receiving dividends (1)	% households selling to cooperative (2)	% actual produce sold to cooperative (3)	% households selling cooperative product to cooperative (4)	Gross value sales to cooperative (12 months) NPR (5)		% households selling cooperative product to other merchants (6)	Gross value sales to other merchants (12 months) NPR (7)		Total values sales NPR (8)
						All	Sellers		All	Sellers	
Baseline	Intervention	31.1	38.6	27.1							
	Comparison	8.8	39.1	20.3							
	Difference	22.2*** (5.4)	-0.5 (7.9)	6.8 (5.0)							
Endline	Intervention	5.3	79.5	62.4	77.3	10,968	14,186	15.2	1,033	11,380	12,001
	Comparison	22.5	72.5	48.3	68.3	11,615	17,518	49.0	4,058	12,075	15,673
	Difference	-17.2*** (6.0)	7.0 (6.7)	14.1** (5.9)	9.0 (7.2)	-646 (2,434)	-3,332 (3,236)	-33.9*** (7.9)	-3,025** (1,413)	-695 (5,266)	-3,671 (3,229)
Change	Intervention	-25.8	40.9	35.3							
	Comparison	13.7	33.4	28.0							
	Difference (PSM diff-in-diff)	-39.4*** (6.7)	7.5 (8.2)	7.3 (7.2)							
OLS with matching weights		-41.2*** (6.8)	7.9 (8.2)	6.4 (7.1)							
# matched intervention observations		132	132	132	132	132	102	132	[illegible]	72	132
# observations total sample		299	299	299	299	299	198	299	[illegible]	20	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

These findings are also confirmed by the results of more subjective measures of the value of sales. Table 9 summarises the answers to a pair of endline questions which asked the respondent about the extent to which their household's income from total sales of the cooperative product (columns 1-3) and other products (columns 4-6) had changed since 2011. Although overall the value of sales seems to have increased in the study area, the responses of the Pavitra members were not significantly different than the responses provided by the comparison households.

Table 9. Subjective perceptions of changes in the value of sales

		Value of sales of cooperative product			Value of sales of other products		
		% households reporting no change (1)	% households reporting increase (less than double) (2)	% households reporting more than double increase (3)	% households reporting no change (4)	% households reporting increase (less than double) (5)	% households reporting more than double increase (6)
Endline	Intervention	26.5	63.6	9.8	34.1	58.3	7.6
	Comparison	24.5	61.2	14.3	25.0	60.8	14.3
	Difference	2.0 (6.4)	2.4 (7.9)	-4.4 (5.9)	9.1 (6.5)	-2.4 (8.1)	-6.7 (6.0)
# matched intervention observations		132	132	132	132	132	132
# observations total sample		299	299	299	299	299	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

5.3. Results on final project outcomes

5.3.1. Livelihoods

Among EDP's key final outcomes was an increase in the incomes of smallholder farmers. Specifically, at the outset of the project, a key objective was for smallholder farmers' income to increase by 300 per cent.

Measuring household's income in non-industrialised countries is not straightforward, particularly in rural areas where respondents tend to be self-employed. Self-reported measures of total income are unreliable, given the wide variety of endeavours such populations engage in to generate income.[9] Moreover, attempts have been made to develop a variety of proxy measures – including data on housing and ownership of fixed assets – but there is currently no consensus on what variables should be included in the analysis.[10] However, given that there is a widely recognised and strong association between household income and consumption,[11] one popular proxy measure employed by the World Bank and other international institutions involves the aggregation of

[9] Morris, Saul, Calogero Carletto, John Hoddinott, and Luc J. M, Christianensen. (1999) *Validity of Rapid Estimates of Household Wealth and Income for Health Surveys in Rural Africa: FCND Discussion Paper No. 72*. Washington: International Food Policy Research Institute.
[10] Hargreaves, James R, Linda Morison, John SS Gear, Julia C Kim, Mzamani B. Makhubele, John DH Porter, Charlotte Watts, and Paul M Pronyk. (2007) "Assessing Household Wealth in Health Studies in Developing Countries: A Comparison of Participatory Wealth Ranking and Survey Techniques from Rural South Africa." *Emerging Themes in Epidemiology*. 4:4.
[11] See Gujarati, Damodar N. (2003) *Basic Econometrics: Fourth Edition*. McGraw Hill, p.4.

household consumption and expenditure data.[12] It is through these data that estimates of the percentages of households living below particular poverty lines are estimated:

Table 10. Livelihoods

		Total per capita expenditures per day (NPR)	Food per capita expenditures per day (NPR)	Poverty £1 (%)	Poverty $2 (%)	Asset index
Baseline	Intervention	137.553	62.408	2.3	16.7	-1.102
	Comparison	139.208	58.766	4.3	11.7	-1.145
	Difference	-1.655 (13.856)	3.642 (3.320)	-2.0 (3.3)	5.0 (5.5)	0.043 (0.126)
Endline	Intervention	219.638	85.207	0.8	8.3	-0.847
	Comparison	255.511	83.724	0.0	2.9	-0.858
	Difference	-35.874* (21.779)	1.483 (5.192)	0.8 (0.8)	5.4 (3.4)	0.010 (0.197)
Change	Intervention	82.084	22.799	-1.5	-8.3	0.255
	Comparison	116.303	24.958	-4.3	-8.7	0.287
	Difference (PSM diff-in-diff)	-34.219 (23.309)	-2.159 (5.011)	2.8 (3.4)	0.4 (5.6)	-0.032 (0.201)
OLS with matching weights		-45.892** (18.506)	2.281 (4.059)	3.6 (3.4)	3.3 (5.4)	-.025 (0.199)
# matched intervention observations		132	132	132	132	132
# observations total sample		299	299	299	299	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

Such data are being employed by the EDP programme to measure its impact on household income. The results are presented in Table 10. Consumption expenditures and poverty lines have been adjusted to account for inflation and changes in exchange rates over the 2011-13 period. The estimates suggest that, although overall, total daily per capita expenditures (column 1) and food per capita expenditures (column 2) have increased over time in the study area (increasing from NPR 138 to NPR 220 for the intervention group and from NPR 139 to NPR 256 for the comparison group), there is not yet evidence of EDP having an impact in this respect. On average, the increase in per capita expenditures was in fact slightly *lower* for Pavitra members (an increase of 60 per cent) compared to their comparison group counterparts (an increase of 86 per cent). However, as noted in Section 4.3, impact was measured only two years after the start of the EDP intervention and the sample excludes households of people who became a member of Pavitra only after the survey visit at baseline. Hence, it is possible that these estimates misrepresent the overall impact of EDP.

Columns 2 and 3 in Table 10 analyse changes in household consumption with reference to two different poverty lines. Note that levels of poverty based on the 1 GBP per day poverty line (column 3) were already quite low at baseline (with two per cent of intervention households considered

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

poor), and therefore allow for little improvement in that respect. There was more scope for impact on the measure of poverty using the 2 USD per day line, with 17 per cent and 12 per cent of intervention and comparison households respectively falling below the poverty line at baseline. We haven't found any evidence of impact of EDP, however, on either of these two measures of poverty.

Finally, column 5 in Table 10 presents statistics for yet another measure commonly used to measure household wealth status, which uses information on ownership of assets and characteristics of the household's dwelling, instead of consumption expenditures. This is based on the assumption that wealthier households are more likely to possess a larger "basket" of locally relevant assets than those which are poorer. This measure is an overall index of these wealth indicators, derived through a statistical method called principal component analysis (PCA). This is a data reduction method that transforms a number of possibly correlated asset variables into a smaller number of uncorrelated variables called principal components. In more simple terms, it narrows in on those assets where there is significant variability in ownership and uses them to assign scores based on the possession of those assets. Hence, households that possess more and less of these assets obtain higher and lower scores respectively. Supporting earlier findings, the asset index indicated no significant difference between the intervention and comparison households in terms of changes in overall asset ownership.

Table 11. Subjective perceptions of well-being

		% Breaking even or doing well (1)	% Struggling (2)	% Breaking even (3)	% Doing well (4)
Baseline	Intervention	64.4	12.4	71.1	16.5
	Comparison	63.1	1.0	87.3	11.7
	Difference	1.3	10.9***	-15.2**	4.3
		(8.1)	(3.6)	(7.0)	(6.1)
Endline	Intervention	93.2	6.8	56.8	36.4
	Comparison	90.8	9.2	72.8	18.0
	Difference	2.4	-2.4	-16.0**	18.4***
		(5.1)	(4.9)	(7.2)	(5.7)
Change	Intervention	28.8	-4.1	-23.7	27.8
	Comparison	27.7	8.6	-16.2	7.6
	Difference	1.1	-12.7*	-4.0	16.7*
	(PSM diff-in-diff)	(9.3)	(7.3)	(11.0)	(9.6)
OLS with matching weights		8.8	-19.9***	-2.8	22.7**
		(9.0)	(7.3)	(10.8)	(9.3)
# matched intervention observations		132	132	132	132
# observations total sample		299	299	299	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

However, evidence based on the more subjective measure of well-being as shown in Table 11 – i.e, the extent to which a household feels able to manage its basic needs, indicated that 20 percentage points fewer intervention households reported themselves to be struggling, whereas 23 percentage points more of these households reported that they were doing well[13]. This suggests that beneficiary households do perceive that they are doing better, or at least have a more positive outlook on life. Note that, in general, subjective measures of well-being risk being biased by psychological patterns. For instance, if respondents *perceive* to have higher total income when payment of sales is done in bulk, rather than in small chunks spread over time, households might *perceive* to be better off now even if total income had remained the same. This is because seeds are sold in bulk at the end of each season, whereas fresh vegetables used to be sold more frequently for smaller amounts each time.

5.3.2. Food security

Table 12. Food security

		% food secure	Total number of meals		Number of main meals		Variety of food types	
			Males	Females	Males	Females	Males	Females
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baseline	Intervention	93.9	3.386	2.962	2.670	2.689	3.830	4.402
	Comparison	100.0	3.207	2.927	2.472	2.775	3.945	4.851
	Difference	-6.1***	0.265	0.035	0.228*	-0.086	0.105	-0.449**
		(2.1)	(0.280)	(0.115)	(0.133)	(0.096)	(0.388)	(0.219)
Endline	Intervention	100.0	4.461	3.750	2.887	2.758	5.261	4.955
	Comparison	99.7	4.313	3.944	2.931	2.914	5.321	5.169
	Difference	3.0	0.156	-0.194	-0.049	-0.156**	-0.067	-0.214
		(4.0)	(0.178)	(0.181)	(0.050)	(0.079)	(0.193)	(0.186)
Change	Intervention	6.1	1.129	0.788	0.196	0.068	1.533	0.553
	Comparison	-3.0	0.887	1.017	0.355	0.139	1.493	0.318
	Difference (PSM diff-in-diff)	6.4	0.181	-0.229	-0.146	-0.071	-0.105	0.235
		(2.2)	(0.383)	(0.207)	(0.135)	(0.124)	(0.534)	(0.263)
OLS with matching weights		7.4***	0.548	-0.257	-0.039	0.007	0.453	0.343
		(2.5)	(0.345)	(0.213)	(0.123)	(0.131)	(0.369)	(0.272)
# matched intervention observations		132	115	132	115	132	115	132
# observations total sample		299	247	299	247	299	247	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

Another key objective of this EDP project was to improve the food security of the targeted population. Both at baseline and at endline, we asked the household respondents in both intervention and comparison groups whether, in the past four weeks, 1) any member in their household had had to eat some foods they really did not want to eat because of lack of resources to

[13] 'Struggling' is defined as 'Managing to meet household needs, but depleting productive assets and/or sometimes receiving support'; 'Doing well' is defined as 'able to meet household needs by your own efforts, and making some extra for stores, savings, and investment.'

obtain preferred types of food; 2) any member had had to eat smaller meals than was needed, or fewer meals in a day, because there was not enough food; 3) there ever was no food to eat of any kind in the household because of lack of resources to get food; and finally 4) any member had to go to sleep at night hungry because there was not enough food. Column 1 in Table 12 reports the results regarding the proportion of food secure households – defined rather broadly as a household for which the answer to *each* of these four questions was 'no'.

Note that, based on this definition of food security, all of the comparison households (that is, 100 per cent) were considered food secure at baseline, leaving zero scope for improvement for this group. By contrast, six per cent% (equivalent to eight households) of intervention households was considered food insecure at baseline. All of these had become food secure by 2011. This improvement in food security could well be the result of the EDP intervention. However, the fact that there was no scope for improvement for the comparison group should be kept in mind when comparing changes over time and interpreting the results.

Columns 2-5 in Table 12 report on an alternative measure of food security, defined by the number of meals consumed per day by the most senior household members, and filtered by gender.[14] Whereas columns 2-3 include snacks, columns 4-5 only include the three main meals. Columns 6-7 present statistics on the variety in the type of foods consumed by the households, ranging from one to eight different types consumed (including grains, roots/tubers, vegetables, fruits, meat, fish, pulses and milk products). Overall, the households in the sample were found to have consumed more meals (snacks) and have diversified their food consumption. However, given the similar trend across intervention and comparison households, we cannot attribute this change to EDP support.

5.3.3. Women's empowerment

Finally, at the outset of the EDP support, the management of Pavitra committed to address the dominance of men in decision-making roles, make it more reflective of Pavitra membership and, in line with the expectations of the project, bring about equal representation of women. It was envisaged that this would be accomplished through leadership capacity development of producer group members and specifically targeting women farmers.

5.3.3.1. Intra-household attitudes towards women

One of the core anticipated outcomes associated with the women's empowerment element of EDP's change model is improved intra-household perceptions of the economic roles of women. To measure this, a set of 15 statements such as *'In general, women are equally capable of contributing to economic well-being as men'* were read out to both male and female respondents of the household survey, and respondents were asked to express the extent to which they agreed or disagreed with each of the statements. Answers ranged from 'strongly disagree' (score 1) to 'strongly agree' (score 6). An overall score out of 90 (i.e. 15 x 6) was computed by taking the sum of the answer scores to all 15 statements. The results are presented in columns 1-2 of Table 13. It is clear that women and men in Pavitra households expressed significantly more positive opinions of women's economic roles in the home and in livelihoods activities. The score from men in particular

[14] Note that some of the senior household members were absent during the survey visit, which explains the smaller sample sizes in columns 2, 4 and 6.

significantly improved: we observed an increase that was almost 10 percentage points higher for Pavitra households than for comparison households.

Table 13. Intra-household attitudes towards women

		Perceptions of women's economic roles Women (1)	Men (2)	Intra-household decision-making score for women (3)
Baseline	Intervention	54.736	52.000	92.807
	Comparison	54.355	49.904	92.637
	Difference	0.555 (1.147)	1.825 (1.635)	-0.560 (1.641)
Endline	Intervention	69.992	68.278	92.795
	Comparison	65.862	59.032	93.557
	Difference	4.130*** (1.009)	9.225*** (1.651)	-0.762 (1.706)
Change	Intervention	15.836	17.132	0.156
	Comparison	11.547	9.466	0.554
	Difference (PSM diff-in-diff)	4.037** (1.754)	8.670*** (3.180)	0.160 (2.434)
OLS with matching weights		5.185*** (1.477)	9.720*** (2.840)	2.903 (2.216)
# matched intervention observations		132	132	132
# observations total sample		299	247	299

Notes: Averages are calculated after matching; Standard errors are shown in parentheses; * significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level.

Another key gender outcome of the EDP programme is increased involvement of women in decisions that affect important life outcomes. Data were captured from the women in relation to this, in particular at the household level. The interviewees were asked to state the extent to which they were involved in 24 areas of household decision-making, ranging from 'no involvement' (score 1) to 'exclusive involvement' (score 5). The areas covered ranged from decisions about what type of foods would be prepared for home consumption to decisions relating to land use planning and management. If the interviewees stated they were not involved or only partly involved in a decision-making area, a follow-up question was asked on the extent to which they could be more involved in this area if they really wanted to be. This was to adjust for the fact that some women may *choose* not to make decisions in particular areas. An overall score out of 120 was computed and the results are summarised in column 3 in Table 13. We have not observed significant changes over time with respect to women's intra-household decision-making scores, with the score remaining close to 93 for both intervention and comparison households.

5.3.3.2. Women's perceptions of their community- and cooperative-influencing capability

The women interviewed were also asked the extent to which they agreed or disagreed with ten positive and negative statements designed to gauge their perceptions of their ability to influence the running of affairs in their communities. Column 1 in Table 14 shows that the increase in the extent to which households reported themselves able to influence community affairs was almost ten points higher for the intervention group than for the comparison group. This difference is statistically significant.

Table 14. Women's perceptions of their community- and cooperative-influencing capability

		Community influence score (1)	Cooperative influence score (2)
Baseline	Intervention	32.530	32.404
	Comparison	35.688	35.687
	Difference	-3.122*** (0.944)	-3.389*** (0.879)
Endline	Intervention	40.712	44.720
	Comparison	36.182	38.048
	Difference	4.530*** (0.848)	6.672*** (1.260)
Change	Intervention	9.017	13.272
	Comparison	0.492	2.348
	Difference (PSM diff-in-diff)	8.282*** (1.408)	10.930*** (1.636)
OLS with matching weights		9.651*** (1.282)	12.028`*** (1.525)
# matched intervention observations		132	132
# observations total sample		299	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

Women who were members of a cooperative were also presented with a series of ten statements to assess their ability to influence the running of affairs in their cooperative. The results, shown in column 2 in Table 14, follow a similar pattern to that observed for community-influencing capability scores: women in intervention cooperatives increased their score by 12 points more than those in comparison cooperatives.

5.3.3.3. Women's self-efficacy

Finally, as discussed in Section 2.1, EDP's change model assumes that women's increased access to and control over income, coupled with improved perceptions of the economic roles of women, will increase their self-efficacy. We used the international self-efficacy scale to measure women's perceptions of their ability to influence issues and events that affect their lives.[15] A self-efficacy score out of 32 was derived by summing up the responses to a set of eight questions, asking the extent to which a particular statement was true, ranging from 'not at all true' (1) to 'exactly true' (4). Statements were of the sort *'you can always manage to solve difficult problems if you try hard enough*', or '*when you are confronted with a problem, you can usually find several solutions*'. The results for all women interviewees are presented in Table 15. Whereas at baseline women in the households of Pavitra members scored similarly to women in comparison households, after the EDP intervention the former group experienced an increase in their self-efficacy score that was almost four points higher than the latter group of women. This difference is statistically significant.

Table 15. Women's self-efficacy

		Women's self-efficacy score
Baseline	Intervention	11.428
	Comparison	11.916
	Difference	-0.489 (0.358)
Endline	Intervention	18.803
	Comparison	15.815
	Difference	2.988*** (0.349)
Change	Intervention	7.375
	Comparison	3.898
	Difference (PSM diff-in-diff)	3.477*** (0.503)
OLS with matching weights		3.670*** (0.534)
# matched intervention observations		132
# observations total sample		299

Notes: Averages are calculated after matching; standard errors are shown in parentheses; * significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

[15] See: http://userpage.fu-berlin.de/health/engscal.htm

6. New Members and Non-member Seed Suppliers: Descriptive Analysis

As explained in Section 2.2.4, one shortcoming of this study is that the data has not allowed us to evaluate the impact of EDP on beneficiaries other than those who were members of Pavitra at baseline (*old members*). In particular, we have been unable to evaluate changes in outcomes of individuals who became members of Pavitra after baseline (*new members*) or people who sell seeds to Pavitra without being a shareholder member (*non-member seed suppliers*). There are various reasons to believe that new members and non-member seed suppliers experienced changes in their lives as a result of the EDP intervention that were different than the changes experienced by older members. For example, it is likely that EDP support allowed Pavitra to expand its membership to groups of people who would never have become a member without EDP's intervention, but whose marginal benefit from membership is relatively high. Consider female- headed households whose livelihoods previously depended entirely on the production and occasional local sale of fresh vegetables. EDP may have supported Pavitra's capacity to reach out to such households and to provide them with services otherwise not available to them, such as access to credit and markets.

To get more insight into the differences in the profile of non-member seed suppliers, new members and old members, we used endline data on a random group of 100 new members and 100 non-member seed suppliers (see Section 4.1) and compared these with data on old members. Table 21–24 in Appendix A3 summarise the results. Asterisks (*) highlight statistically significant differences between the value of new and non-members and the equivalent value of old members. A few results stand out.

Unsurprisingly, we found that households of the older Pavitra members were relatively large (six members compared to five), had more adult members (four compared to three) and were headed by an older person than households of new members. Households of non-member seed suppliers were similar to new members in terms of their demographic characteristics, except for the average age of the head of the household, which seemed to be closer to that of old members. Interestingly, compared to old members, new member and non-member seed supplying households were significantly more likely to be headed by a female person. Although this result could be partially driven by the demographic differences between old and new members, it would also be consistent with Pavitra having increased its efforts to reach out to female-headed households since EDP's intervention.

As we would have expected, non-members reported having significantly lower access to services provided by the cooperative. Although we did not observe any differences regarding their access to advice and technical support in the field and provision of agricultural inputs (such as foundation seeds), they were less likely to have received training in improved agricultural methods and less likely to have accessed production loans and PLC classes. They were also less likely to have received fertilisers from Pavitra, which is consistent with qualitative data suggesting that Pavitra gives priority to its own members in case of supply shortage. However, it is less clear why new members were ten percentage points less likely to have received fertilisers over the previous 12 months than old members. Despite non-members being less likely to have accessed services provided by Pavitra, those who did make use of the training in improved agricultural methods and the provision of fertilisers seemed to have been significantly more pleased with the quality of these services.

Compared to old and non-members, new members were relatively happier with the quality of the PLC classes organised by Pavitra.

An important finding of the core analysis in Section 5 was that Pavitra members in our sample (i.e. old members) did not change their production towards commercial crops/seeds any more than their comparative counterparts, nor did they decrease their use of inorganic fertilisers. However, Table 23 in Appendix A3 suggests that this result might have been different for new Pavitra members, who at endline were more likely to have reported producing commercial crops or seeds than old members and less likely to have used inorganic fertilisers on their farms. It is possible that older members, given their age and experience, are more reluctant to change their old practices and hence might have been less likely to benefit from the advice and support provided by EDP. This suggests that for new members the overall impact of EDP might have been larger than the effects we have estimated above. The opposite may be true for non-member seed suppliers, who seem to have been less likely to have engaged in compost-making activities, such as use of animal urine or improved animal manure.

A similar pattern can be observed for access to loans. Compared to old members, a higher proportion of new members reported to have obtained a loan from Pavitra, whereas non-members were less likely to have accessed a loan. Again, this is consistent with new members potentially being more dynamic than older members and non-members having less access to cooperative services. Households of new members and non-member seed suppliers had on average fewer male and female members engaged in agriculture. Again, this could be a result of differences in demographic characteristics. Compared to old members, they were also less likely to have received dividends. Importantly, in comparison to the households of old- and non-Pavitra members, the female members in the households of new Pavitra members made significantly more contributions to savings associations, such as the saving scheme run by Pavitra.

However, an unexpected result is that only four percent of the 100 new members sampled reported to be a member of a farmer group, whereas this was 58 and 92 per cent for old members and non-member seed suppliers respectively. This is rather disturbing, since EDP encouraged Pavitra to make farmer group membership a requirement of their cooperative membership eligibility criteria. This result calls for further urgent investigation.

Non-member seed suppliers reported having sold a higher proportion of their produce to Pavitra. This is not surprising since, per the sampling strategy, this group is biased towards households that actually sold their produce to Pavitra. More interesting is the finding that, relative to older members, a significantly higher proportion of new Pavitra members (by approximately 15 percentage points) reported experiencing an increase in the value of their sales to Pavitra and other merchants, whereas a lower proportion reported having experienced no change. This result is consistent with the finding of a higher average value for total consumption expenditures and food consumption expenditures at endline for new members, and higher scores on all of the women's empowerment indicators, except for intra-household decision-making power. These descriptive statistics re-confirm the importance of the aforementioned caveat that the estimated impacts we measured above may be underestimating the overall impact of EDP.

7. Conclusions

7.1. General findings

The main objective of EDP's intervention in Pavitra, Nepal, was to improve the livelihoods and the food security of the cooperative's members and to strengthen the empowerment of women in particular, both in the enterprise and in their respective communities and households.

The most promising findings of this study relate to EDP's impact on a range of indicators of women's empowerment. Both women and men in Pavitra households expressed significantly more positive opinions of women's economic roles in the home and in livelihoods activities than in the comparison group. Significantly more women agreed with positive statements about their ability to influence affairs in their community, and even more so in the cooperative. Female members of Pavitra households increased their level of self-efficacy (that is, their perceptions of their ability to influence issues and events that affect their lives) to a greater extent than the comparison women. We have not observed any significant changes over time with respect to women's power to influence decisions within their household.

Although we have observed improved livelihoods for the beneficiary households, when assessed through an evaluation of household consumption levels, we have also seen similar trends for the matched comparison group and therefore conclude that there is no evidence yet that the EDP project is having an impact on material well-being. We found no significant changes in the proportion of households falling below the various poverty lines, the level of food security measured by the number of meals consumed (differentiated by gender) and the variety of food consumed. We should note however that, given that the households in this area were already quite food secure at baseline and since only a few of them fell below the standard poverty lines, there was little scope for improvement in this respect. There was also no evidence of impact when looking at improvements in indicators of material wealth such as quality of housing or ownership of assets. By contrast, using a more subjective measure of well-being, Pavitra member households were significantly less likely to describe themselves as struggling to meet their basic needs, and more likely to describe themselves as doing well. Although subjective evidence of this sort needs to be treated with caution, it does suggest that EDP has enabled Pavitra members to gain a more positive outlook on life.

A few important qualifications need to be made when interpreting the results on 'final outcomes'. First, impacts have been examined only two years after the start of the EDP intervention, and should thus be read as interim rather than final. Second, given the huge expansion of Pavitra membership and the number of non-member seed suppliers in the two and a half years since EDP started, the study sample drawn at baseline does not represent the overall impact of the project. Studying endline characteristics of a sample of 100 new members, we have observed some significant differences with the study sample of older members (that is, those who joined Pavitra before the EDP intervention). In particular, we recorded a significantly higher average value for total consumption expenditures and food consumption expenditures and even higher scores on most of the women's empowerment indicators. In combination with a few other caveats highlighted in this report, these facts imply that that the improvement in outcomes such as material well-being observed on average in this study may be an understatement of the true benefit derived from EDP by the bulk of farmers. In order to understand how these improvements (or absence of improvements) came about as evidenced in final outcomes, this study also evaluated Pavitra's performance against a set of intermediate outcomes. Specifically, we looked at indicators related to access to quality services, the adoption of improved agricultural practices, use of land and labour in agriculture, crop yields and households' revenues from the cooperative (through dividend payouts or sales).

Pavitra members were more likely than comparison cooperative members to report that their cooperative was providing fertilisers, production loans and women's participatory classes and that the quality of advice and technical support and the provision of agricultural inputs (such as foundation seeds) had improved. However, they were no more likely to report having *used* these services. In particular, access to production loans remains restricted to relatively few members, with only 24 per cent of the households in the study sample having reported that they took out a loan over the 12 months prior to the survey. It is also worth noting that new members were ten percentage points more likely than old members to have received a loan. However, the fact that 66 per cent of households did not access loans suggests that some barriers to loan uptake remain (either on the demand or the supply side). We have not found any evidence of impact on the savings behaviour of either male or female members of the households in the sample.

To some extent, the improved quality of advice and support in agricultural practices seems to have paid off. Pavitra members are now more likely to undertake their own compost-making activities and to irrigate their fields. However, compared with the comparison group, we have not observed differential changes in the use of animal urine and improved animal manure. The use of inorganic (chemical) fertilisers has increased more rapidly for Pavitra members than for their comparison group counterparts. We have not observed any significant differences in the type of crops or seeds cultivated by the farmers – that is, both comparison and intervention groups reported to be equally commercial in the type of products that they are cultivating. However, looking again at the endline data on our sample of 100 new members, we can see that new members are significantly more likely than old members to have produced commercial crops or seeds and less likely to have used inorganic fertilisers on their farms. This evidence suggests that older members, perhaps due to their age and extensive experience, may be more reluctant to change their practices, and hence might be less likely to benefit from the advice and support provided by EDP. This could explain the lack of impact on the adoption of improved agricultural practices and on overall average household incomes observed for the members in our study sample.

We have found no significant changes over time in terms of the area of land owned, the area of land used for farming, or the area of land allocated to the production of the cooperative product. While there were no changes in the number of household members working on the land, the quantity of *male* labour hired from *outside* the household did increase. However, since a similar upward trend was observed in the comparison group we cannot attribute this change to EDP's intervention. Although we haven't observed changes in the area of land cultivated, a higher proportion of Pavitra members reported to have experienced increased yields on their farms over the last 12 years. This effect is not statistically significant, however, and should therefore be interpreted with caution.

Surprisingly, we have not observed evidence of the EDP's impact on farmer group membership. Most strikingly, only four out of 100 households that have joined Pavitra since 2011 reported being a member of a farmer group at the time of the endline survey. Given that EDP explicitly encouraged

Pavitra to include farmer group membership in their cooperative membership eligibility criteria; this is a disconcerting result which requires further investigation.

Pavitra members have been significantly less likely than comparison groups to receive dividends over the last year. This result could be explained by the fact that Pavitra is in the midst of investing in the construction of a processing plant and therefore its profits are currently relatively low.

Overall, we did not find evidence that EDP has enabled the beneficiary group to increase the quantity or the value of its sales to the cooperative to a level that would have been impossible without EDP support. These findings are supported by the results from more subjective measures that asked the respondents about the extent to which the value of their sales had increased since EDP's intervention. The proportion of households that reported making increased sales was approximately the same as in the intervention and comparison groups.

7.2. Learning considerations

Based on the evaluation results, we make the following set of key programme learning recommendations:

- ***Continue to build on the significant successes in women's empowerment***
  The results clearly show that the EDP intervention has successfully supported Pavitra in improving women's self-efficacy and their say in the running of the affairs of their enterprises and broader communities. The next step will be to find creative ways to strengthen women's influence at the household level, which – as the evaluation shows –- remains limited to date.

- ***Monitor impact on livelihoods***
  The evaluation does not yet find an impact on the livelihoods of the farmers who were members of Pavitra at baseline. Although we have observed significant improvements in the value of their sales and their consumption expenditures, this positive trend is similarly present for the members of the comparison cooperatives in the area. If EDP's aim is to lift Pavitra members' livelihoods to a higher level – that is, a level that would not have been possible without EDP's support – it may be worth investing resources into more innovative ways to boost incomes.

- ***Investigate the reason for Pavitra members' low participation in farmer groups***
  We have not observed evidence of EDP's impact on farmer group membership. This result could partially be driven by the fact that baseline data were collected with a delay of six months. However, only 58 per cent of Pavitra's older members and only four per cent of Pavitra's newer members reported being members of a farmer group by the time of the endline survey. Given that EDP explicitly encouraged Pavitra to only consider farmers who were members of a producer group as eligible cooperative members, this result is rather disconcerting and requires further inquiry.

- ***Monitor dividend payouts to ensure all shareholders benefit from Pavitra's increase in profits***
  The evaluation reveals that almost none of the Pavitra shareholders received any dividends over the last year. In fact, a significantly smaller proportion received dividends compared to the period before EDP's intervention. This finding can be convincingly explained by the fact that profits have been invested in the construction of a processing plant. However, it is important to monitor the pay out of dividends in the coming years to make sure all shareholders benefit from Pavitra's increase in profits.

- ***Consider further research to investigate why farmers' access to loans remains limited***
  This study has the disconcerting finding that loan uptake remains limited to relatively few Pavitra members, i.e. 21 per cent. Older Pavitra members are least likely to obtain loans. It is unclear what is driving this result. One reason might be the interest costs. At the outset, it was expected that Pavitra would be able to provide loans at favourable interest rates. The evaluation reveals, however, that the interest rate of 16 per cent offered by Pavitra is actually higher than the average interest rate offered by comparison cooperatives (15 per cent). Another reason might be Pavitra's recent policy change regarding loan eligibility. Qualitative evidence gathered in the context of this review hints that Pavitra has changed its policy, so that only members who own a minimum amount of land *and* who are able to present a land certificate are currently eligible for loans from Pavitra. Not only does such a policy deny access to loans to landless farmers, it could also form a loan barrier to women, especially given that their intra-household bargaining power remains limited. This result is worth further investigation.

- ***Continue mentoring on improved agricultural practices***
  To some extent, the improved quality of advice and support in agricultural practices has paid off as farmers have become more likely to engage in compost-making activities and irrigation of their fields. EDP may want to consider alternative ways to encourage farmers, especially the older Pavitra members, to substitute inorganic fertilisers with organic fertilisers such as animal urine and improved animal manure.

- ***Continue to strengthen and build on the evaluation process***
  EDP should be commended on the efforts it has invested in designing and carrying out a rigorous impact evaluation. In many ways, the evaluation process adhered to in this EDP project is a considerable improvement compared to other evaluations conducted in the sector. However, as always there is scope for improvement to the survey instruments and sampling strategy. Section 4.3 in this report lists a set of caveats, mainly related to data availability. We strongly recommend that EDP carefully considers these shortcomings in order to strengthen the evaluation process for future EDP projects.

Appendix A1. Details about the Choice of Evaluation Strategy

Limitations in pursuing the gold standard

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

Intervention effect = average post-programme outcome of participants minus what the average post-programme outcome of these same participants would have been had they never participated.

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

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

Unfortunately – outside the context of specially designed pilot studies – randomised evaluation designs are seldom implemented in the context of social programmes run by international NGOs such as Oxfam. There can be cost, feasibility, and/or ethical constraints that militate against their use or simply due to the desire among implementing agencies to work with purposely chosen populations. Moreover, there are often cases where the *opportunity* to participate in a programme is put in place – as would be the case with the setting up of a micro-credit programme – and people *choose* whether to participate. Those that choose to participate are likely to be different than those that do not, including in characteristics that are intrinsically difficult to measure, e.g. motivation.

The EDP Pavitra project in Nepal forms a clear example of a case where the selection of beneficiaries was not done randomly. As discussed in Section 2.2.1, the producer group being targeted was purposely selected, with Pavitra being identified as an eligible cooperative through a careful consultation process, and individual participants, in turn, being selected based on Pavitra's membership eligibility criteria. As such, participation was not random and the evaluation team had to identify appropriate, non-random comparison groups for the purpose of this study. The selection of the comparison group will be discussed in more detail below.

Despite the fact that the comparison group is carefully selected to be as similar as possible to the intervention group (and given that the comparison group is non-random, and taking into account known differences such as the sector), the two groups may differ in a) their pre-programme outcomes; and b) their other observed and unobserved characteristics. As such, directly comparing the outcome differences between the intervention and comparison groups may well lead to biased intervention estimates.

Alternative evaluation design: matched difference-in-difference

Fortunately, there are several evaluation designs that can be used when the comparison group is non-equivalent that can – particularly when certain assumptions are made – generate reasonably precise intervention effect estimates. One is the difference-in-difference method. Here, the comparison group is not used, in and of itself, to estimate the counterfactual. Rather, the *trend* in outcome-related change experienced by the comparison group is used to estimate what this trend would have been for the intervention group in absence of the intervention. While the intervention and comparison groups are allowed to have different potential outcomes in absence of the intervention, a “parallel trend” assumption is made, i.e. the *extent* of outcome-related change experienced by the two groups would be *the same* in a counterfactual scenario. The intervention effect is calculated by subtracting outcome gains made in the intervention and comparison groups. If there is no positive and statistically significant difference among the differences, then it is assumed that the intervention had no effect on the outcome variable in question and vice-versa if a positive and statistically significant difference is observed.

Table 16 provides a simple example of how the difference-in-difference intervention effect estimate is calculated. First observe how ignoring the trend experienced by the comparison group and only considering the change in outcome of the intervention group would yield a biased estimated increase of £50. Similarly, notice how a direct follow-up comparison of the intervention and comparison group would yield a significantly higher – and hence upwards-biased – estimate, i.e. £60. The correct estimate in this table would be the difference-in-difference estimate of £40.

Table 16. Difference-in-difference intervention effect estimate – example

	Mean baseline income	Mean follow-up income	Difference between time periods
Intervention group	£50	£100	£50
Comparison group	£30	£40	£10
Difference between groups	£20	£60	£40

Another method popularly used in non-randomised programme evaluation is matching. As mentioned above, one of the inherent problems with non-randomised comparison groups is that they often differ in ways that influence the potential outcomes of programme participants and, consequently, do not serve as good counterfactuals. One solution to this problem is offered by matching: find units in the comparison group that possess the same baseline characteristics – e.g. ethnicity, age and sex – as those of the intervention group and match them on these characteristics. If matching is done properly in this way, the observed characteristics of the matched comparison group will be identical to those of the intervention group. The problem, however, is that with larger and larger numbers of characteristics on which to match, it becomes more and more difficult to find comparison units with similar combinations of characteristics for each of the units in the intervention group. The end result, typically, is that only a few units from the intervention and comparison groups get matched up, thereby not only significantly reducing the size of the sample but also limiting the extent to which the findings can be generalised to all programme participants. This is referred to as the "curse of dimensionality" in the literature.

Fortunately, matching on the basis of the propensity score – i.e. the conditional probability of being assigned to the programme group, given particular background variables or observed characteristics – offers a way out. The way propensity score matching (PSM) works is a follows: units from both the intervention and comparison groups are pooled together. A statistical probability model is estimated, typically through logit or probit regression. This is used to estimate programme participation probabilities for all units in the pooled sample. Intervention and comparison units are then matched within certain ranges of their conditional probability scores. Tests are further carried out to assess whether the distributions of baseline characteristics are similar in both groups after matching. Provided that a) the dataset in question is rich and of good quality; b) the groups possess many units with common characteristics (i.e. there is a large *area of common support*); and c) there are no unobserved differences lurking among the groups, particularly those associated with the outcomes of interest, PSM on its own can produce good intervention effect estimates.

The validity of both the difference-in-difference and PSM designs are heavily dependent on their underlying assumptions. The parallel trend assumption associated with the former method, for instance, can be called into question if there are differences existing between intervention and comparison units (e.g. educational status) that differentially interact with changes in the external environment (e.g. a new government programme that favours literate people), thereby resulting in differential trends in outcome change, even in the absence of the intervention. The validity of PSM, on the other hand, is founded heavily on the "selection on observables" assumption, and therefore treatment effect estimates can be biased if there are unmeasured but relevant differences existing between the groups.

Fortunately, the robustness of both the difference-in-difference and PSM methods can be strengthened when they are used together. With matching on observable characteristics via PSM, the parallel trend assumption associated with the difference-in-difference method is only required for the unmeasured differences between the groups, given that PSM ensures that the two groups are observationally the same. In other words, it is only the influence of the participants' *unobservable characteristics* on outcomes that needs to remain the same over time. Intervention effect estimates are also no longer heavily dependent on the "selection on observables" assumption, given that any relevant unobserved differences between the groups – provided the parallel trend assumption for unobserved differences holds – get cancelled out via differencing. EDP Nepal Pavitra's evaluation design is based on this integration of the difference-in-difference and PSM methods. In addition, we will use Ordinary Least Squares (OLS) regression techniques to test the robustness of our results to the controlling for changes in observable characteristics over time.

Choice of non-randomised comparison group

A key factor in ensuring the validity of any non-randomised impact evaluation design is the use of an appropriate comparison group. Comparison groups which differ in crucial unobserved time-varying baseline characteristics which cannot be controlled for by use of the double difference approach and/or which are subjected to different external events and influences, will likely result in misleading conclusions about programme impact. Identifying a plausible comparison group is therefore critically important and is, generally speaking, not an easy task in non-experimental work.

Ideally, we would have identified cooperatives in neighbouring VDCs that were active in the same sector, i.e. vegetable seeds, as cooperatives from which to draw a sample of comparison members for our study. As expected, however, locating a cooperative from the vegetable seeds sector within the same geographical area was found to be problematic. Instead, we used the results of the consultation process presented in Figure 2 to identify agricultural cooperatives in neighbouring VDCs that were active in sectors most similar to vegetable seeds in terms of EDP's top-level targeting criteria, i.e. activities that 1) benefit women and smallholders and 2) have high income potential. Dairy farming, fresh vegetables and spices were identified as activities that came closest to vegetable seeds in their ranking based on these characteristics (see Figure 2).

Within these sectors, we identified three comparison cooperatives located in the neighbourhood of the intervention area – but in different VDCs: Bheri Krishi (cereal seeds and fresh vegetables), Bheri Pashu Tatha Krishi (ginger, goat rearing and grass production) and Bhaktapur (cereal seeds). The members of these cooperatives live in five neighbouring VDCs, namely Sahare, Ghumi, Malarani, Ghumkhahare and Rhamgat. After baseline data were collected, Pavitra expanded its membership to two additional VDCs, namely Sahare and Ghumkhahare. Whereas the membership expansion did not affect the comparison group in Sahare (the membership only reached one village in this VDC which was sufficiently far away from the comparison households to be affected by the intervention), the comparison households in Ghumkhahare had become part of the treatment group and a decision was made to drop these households (only four in total) from the study. This yielded a study population with households from three intervention VDCs and four comparison VDCs, the location of which is indicated in Figure 5.

Some identifying conditions needed to be satisfied for our comparison group – and hence our evaluation design – to be valid. First, we needed to ensure that the comparison group in no way benefitted from the intervention, directly or indirectly. According to the Cooperative Act 9092 and the Cooperative Act 92, a cooperative in Nepal can provide its services to a maximum of only five Village Development Committees (VDCs).[16] These regulations favoured our evaluation design in that

[16] It should be noted that recently, there has been a change in regulations, allowing cooperatives to expand some of their services to other VDCs, conditional on permission from the Divisional Cooperative Office (DCO)

they ensured that the comparison groups, located in different VDCs, did not have any direct access to any of the services provided by Pavitra.

More difficult to reject is the possibility that the comparison households benefitted from *spill-over effects*, i.e. with the impact potentially being passed on from direct beneficiaries to comparison households, for instance through transfer of knowledge on the use of improved agricultural methods. Farmer-to-farmer transfer of information across VDCs was expected to be minimal, given that VDCs – each consisting of multiple villages – are quite big and difficult to reach for farmers in this area who typically travel on foot. However, it is quite likely that cooperative representatives learn from each other and provide each other with advice and support, which in turn may indirectly benefit their members. Although we did not expect such indirect effects to be substantial, this caveat does imply that by comparing the outcomes of the comparison and intervention households we may have been underestimating the true extent to which EDP has impacted on Pavitra members. One may also be concerned that EDP's intervention in Pavitra may have negatively affected other cooperatives in the area, by making Pavitra more competitive. During qualitative interviews, representatives of both intervention and comparison cooperatives firmly rejected this concern with the argument that Pavitra specialises in different types of goods, targeting very different and independent markets.

Finally, we needed to be certain that the intervention and comparison households were similarly affected by any external factors other than the EDP intervention, such as climate shocks and support from other NGO/governmental actors, in terms of the outcomes of interest in this study (e.g. income levels). The selection of comparison VDCs located in the near neighbourhood of intervention VDCs largely minimised the issue of differential climate shocks. Furthermore, qualitative interviews with NGO and government representatives of intervention and comparison VDCs suggested that, although all cooperatives received support from other organisations – especially before the EDP intervention – the support related to the outcomes of interest in this study had been fairly similar across the board.

Technical details of the PSM procedure employed in this study

As discussed above, our impact assessment analysis combines the propensity score matching method (PSM) with the difference-in-difference estimator, to control for both observed baseline differences between intervention and comparison groups (through matching on such characteristics) and the effect of unobserved baseline differences that are constant over time (by taking the difference in *changes* in outcomes over time, rather than outcome *levels*). In short, this was achieved by first matching intervention households with those comparison households that were observationally similar to them at baseline (using a particular matching algorithm) and then, in a next step, by comparing the changes in the outcomes of the intervention households with the equivalent changes in outcomes of their respective matches. Unsurprisingly, there are different approaches to matching, i.e. to determining whether or not a household is observationally 'similar' to another household. For an overview, we refer to Caliendo and Kopeinig (2008).[17] The following section describes the matching procedure followed in this study.

in Surkhet district. At the time of the follow up data collection visit, however, this recent change had not yet affected the comparison groups of our study.

[17] Caliendo, M. And Kopeinig, S. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching", *Journal of Economic Surveys*, Wiley Blackwell, vol. 22(1), pages 31-72.

Estimating propensity scores

First, we pooled the units from both the intervention and comparison groups together and used a statistical probability model (i.e. a *probit* regression) to estimate *propensity scores* for all units in the pooled sample, i.e. the probability of being a member of Pavitra, conditional on a set of observed characteristics. Tables 17 and Table 18 present the probit regression results used to estimate the propensity scores on the basis of which the matching was subsequently done. Table 17 shows the probit results for the non-parsimonious model entering the full set of matching variables considered in this study. The specific choice of matching covariates was guided by discussions with the EDP programme team and Pavitra cooperative representatives regarding 1) the criteria used in the selection of Pavitra as a cooperative eligible for EDP support and 2) the cooperative's membership eligibility criteria. To guarantee that none of the matching variables were affected by the intervention, we only considered variables that were measured at baseline, and only those variables that were unlikely to have been influenced by anticipation of the EDP intervention.

Table 17. Estimating the propensity score – non-parsimonious model

Dependent variable: Pavitra membership = 1			
Explanatory variables/characteristics	Coefficient		p-value
Household size	-0.046	(0.043)	0.282
Head of household in good or fair health = 1	0.097	(0.422)	0.818
Proportion of children < 12 years old	-0.289	(0.469)	0.538
Head of household is female = 1	-0.027	(0.305)	0.929
Head of household is > 60 years old = 1	-0.385	(0.250)	0.124
Head of household received secondary education = 1	-0.255	(0.186)	0.171
Area of land used for farming (in hectares)	1.446	(0.635)	0.023
Area of land owned by household	-1.163	(0.608)	0.056
No. of months household has been member of producer group	-0.001	(0.002)	0.584
Household involved in agricultural activities = 1	-0.726	(0.265)	0.006
Household involved in income generating activities = 1	0.058	(0.225)	0.798
Household involved in casual labour = 1	0.084	(0.229)	0.713
Household involved in full time work = 1	0.353	(0.201)	0.079
Wealth index	-0.381	(0.051)	0.000
Household's livelihood depends for more than 50% on agriculture = 1	0.247	(0.173)	0.153
Number of observations	319		

Notes: Probit regression. All characteristics are measured at baseline in 2011. Standard errors are shown in parentheses. Explanatory variables expressed as x = 1 represent binary variables taking values of either 0 or 1. The dependent variable is 1 if the household is a member of the Pavitra cooperative and 0 otherwise. The coefficients represent the contribution of each explanatory variable/characteristic to the probability that a household is a member of Pavitra.

Note that the probability estimation equation is not a determinants model. For this stage of PSM, causality is not of as much interest as the correlation of the covariates with treatment. We should therefore refrain from over-interpreting these results. Note that, on average at baseline and relative to the comparison group, fewer Pavitra households reported being involved in agricultural activities and more of them claimed to be involved in full-time salaried work. Although they owned less land, they did report having a larger area of land devoted to cultivation. Most importantly, they were much poorer according to the asset-based index and this is an important feature to take into account for the purpose of selection bias correction.

Table 18, in turn, shows the results of the probit model restricted to the sub-set of eight variables which were correlated with being a member of Pavitra at *p*-values of 0.20 or less, determined through backwards stepwise regression.

Table 18. Estimating the propensity score – Restricted model

Dependent variable: Pavitra membership = 1			
Explanatory variables/characteristics	Coefficient		p-value
Head of household is > 60 years old = 1	-0.438	(0.227)	0.053
Head of household received secondary education = 1	-0.250	(0.179)	0.162
Area of land used for farming (in hectares)	1.418	(0.637)	0.026
Area of land owned by household	-1.231	(0.607)	0.043
Household involved in agricultural activities = 1	-0.765	(0.244)	0.002
Household involved in full time work = 1	0.308	(0.194)	0.113
Wealth index	-0.380	(0.049)	0.000
Household's livelihood depends for more than 50% on agriculture = 1	0.233	(0.167)	0.162
Number of observations	319		

Notes: Probit regression. All characteristics are measured at baseline in 2011. Standard errors are shown in parentheses. Explanatory variables expressed as x = 1 represent binary variables taking values of either 0 or 1. The dependent variable is 1 if the household is a member of the Pavitra cooperative and 0 otherwise. The coefficients represent the contribution of each explanatory variable/characteristic to the probability that a household is a member of Pavitra.

In the main text we used the propensity score estimates of the latter model to match households. It should be noted, however, that the results remained substantially similar when we used the results of the non-parsimonious model as a basis for matching (results not reported in this review).

Defining the region of common support and balancing tests

Intervention and comparison units were then matched on the basis of the estimated propensity score, using the kernel method (through the *psmatch2* module in STATA using the default of 0.06 as a bandwidth) on the area of common support. The area of common support is the region where the propensity score distributions of the treatment and comparison groups overlap, which in our case turns out to be quite large, i.e. [0.04575757, 0.96890045]. All of the 132 treated households in our sample could be matched using this procedure, suggesting that the careful choice of our comparison groups at baseline has been successful. We only lost 18 comparison households by restricting the sample to those households that fall within the area of common support.

Furthermore, for PSM to be valid, the intervention group and the matched comparison group needed to be balanced in that they needed to be similar in terms of their observed characteristics. This should be checked. The most straightforward method to do this is to test whether there are any statistically significant differences in covariates between the treated and control group in the matched sample. In Table 19, balance was checked on all available covariates listed in Table 17, even if they were not included in the final estimation shown in Table 18. We confirm that the covariates were balanced across groups at *p*-values greater than 0.20.

Table 19. Balancing test

Baseline characteristic		Mean		p-value
		Treated	Control	
Household size	Unmatched	5.492	5.894	0.100
	Matched	5.492	5.763	0.297
Head of household in good or fair health = 1	Unmatched	0.947	0.941	0.820
	Matched	0.947	0.948	0.964
Proportion of children < 12 years old	Unmatched	0.216	0.214	0.996
	Matched	0.216	0.206	0.683
Head of household is female = 1	Unmatched	0.091	0.083	0.805
	Matched	0.091	0.088	0.934
Head of household is > 60 years old = 1	Unmatched	0.144	0.183	0.363
	Matched	0.144	0.142	0.971
Head of household received secondary education = 1	Unmatched	0.326	0.432	0.061
	Matched	0.326	0.324	0.981
Area of land used for farming (in hectares)	Unmatched	0.469	0.493	0.480
	Matched	0.469	0.450	0.564
Area of land owned by household	Unmatched	0.423	0.480	0.101
	Matched	0.423	0.419	0.913
No. of months household has been member of producer group	Unmatched	23.614	27.787	0.315
	Matched	23.614	26.936	0.439
Household involved in agricultural activities = 1	Unmatched	0.803	0.923	0.002
	Matched	0.803	0.798	0.925
Household involved in income generating activities = 1	Unmatched	0.174	0.160	0.739
	Matched	0.174	0.142	0.480
Household involved in casual labour = 1	Unmatched	0.242	0.136	0.018
	Matched	0.242	0.202	0.431
Household involved in full time work = 1	Unmatched	0.273	0.314	0.442
	Matched	0.273	0.229	0.424
Wealth index	Unmatched	-1.1024	0.483	0.000
	Matched	-1.1024	-1.147	0.824
Household's livelihood depends for more than 50% on agriculture = 1	Unmatched	0.432	0.414	0.760
	Matched	0.432	0.459	0.661
Average standardised percentage bias[1]	Unmatched	18.05		
	Matched	5.02		
Number of observations	Unmatched	319		
	Matched	301		

Notes: The average standardised percentage bias is the % difference of the sample means in the treated and control sub-samples as a percentage of the square root of the average of the sample variances in the treated and control groups. As a rule of thumb, a bias of 5 or less is tolerated.

Matching intervention households to comparison households

Finally, our choice of matching technique using kernel matching was driven by the consideration of the relatively small sample size and the associated need to use as much information in the data as possible. The risk with alternative methods, such as nearest neighbour matching, is that only a small number of comparison households is captured in the area of common support. Matching estimators such as kernel matching instead use a weighted average of all comparison units to construct the counterfactual match for each intervention household.

Finally, standard errors were bootstrapped (stratified by cooperative) using 1000 repetitions to account for the additional variation caused by the estimation of the propensity scores and the determination of the common support.[18]

Appendix A2. Detailed Guide to the Interpretation of Results

Impact or effect size

Simply put, impact refers to the size of the difference between groups when evaluating outcomes. In this report, impact will often be stated as the average difference between people who received Oxfam support (intervention) and matched people who did not (matched comparison).

Statistical significance

When we talk about 'significant impact' in this report, we mean impact that is 'statistically significant'. For example, let's say we measure that the average sampled Oxfam beneficiary household earned NPR 50,000 more than the average sampled comparison household. Although this difference seems quite large, we remember that this estimated average impact is derived from data on a *sample* of intervention and comparison individuals, rather than data on the *true* population of beneficiaries and their respective counterfactual values (see discussion in Appendix A1. Given that we only have one sample available, there is always the off-chance that we have been unfortunate with our sample and that we happen to have drawn beneficiaries with relatively high values from a population pool with a true average impact of zero. Therefore, it is necessary to take into consideration the statistical probability of measuring impacts of NPR 50,000 if there were truly no differences in earnings between the intervention and comparison groups, i.e. if in reality there were zero impacts. This probability is usually referred to as the *p-value*. P-values help to evaluate study hypotheses. The default hypothesis is always that there are no differences between the intervention and comparison groups. When a difference is detected, the p-value is used to evaluate whether the default hypothesis should be rejected – i.e. that there are no differences between the groups – and conclude that the project had an impact. If the p-value is small, for instance one per cent, this means that the probability of obtaining an estimate of NPR 50,000 if the true impact was zero is very small and we can reject the hypothesis that the project had no impact on earnings at the one per cent significance level. We would then say that the

[18] Bootstrapping is a statistical procedure where repeated samples are drawn from the original sample and parameters, such as standard errors, and re-estimated for each draw. The bootstrapped parameter is calculated as the average estimate over the total number of repeated draws.

result is 'significant'. This is different than saying that the magnitude of the result is large, because magnitude refers to effect size or impact, as explained above. Taken together, significance and magnitude help us to interpret the results.

In the results tables below, statistical significance will be indicated with asterisks, with one asterisk (*) indicating a p-value of less than one per cent, two asterisks (**) indicating a p-value of less than five per cent and three asterisks (***) indicating a p-value of less than ten per cent. The higher the p-value, the less confident we are that the measured estimate reflects the true impact. Values with a p-value of more than ten per cent are considered statistically insignificant in this report. Note that the smaller the sample size and the greater the variation in the outcomes (i.e. *standard errors*) among the sampled households, the less confident we are in the accuracy of the estimated impact, the larger the p-value, and, hence, the less likely we are to conclude that a result is statistically significant.

Fictitious example

To illustrate, consider Table 20, which presents a fictitious example of the type of results tables that are used in the results sections in this report.

Table 20. Fictitious data to illustrate the interpretation of results

		Outcome X (NPR) (1)	Outcome Y (NPR) (2)	Outcome Z (NPR) (3)
Baseline	Intervention	150	150	150
	Comparison	50	148	148
	Difference	100***	2	2
		(0.5)	(0.5)	(0.5)
Endline	Intervention	150	300	300
	Comparison	50	296	150
	Difference	100***	4	150***
		(0.6)	(0.6)	(0.6)
Change	Intervention	0	150	150
	Comparison	0	148	2
	Difference	2	2	148***
	(PSM diff-in-diff)	(0.5)	(0.5)	(0.5)
OLS with matching weights		0.5	3	147***
		(0.4)	(0.4)	(0.4)
# matched intervention observations		132	132	132
# observations total sample		299	299	299

Notes: Averages are calculated after matching; standard errors are shown in parentheses;
* significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.

Assume we are interested in the impact of a project on three outcomes X, Y and Z. Table 20 presents the results using our sample of 299 matched households, out of which 132 are intervention households (i.e. Pavitra members) and the remaining 167 households are members of the comparison cooperatives. For both matched intervention and comparison groups, the table reports the average values observed at two points in time, i.e. at *baseline* and at *endline*. We also present the observed *changes* in the outcomes between baseline and endline. In line with the discussion above, statistical significance of differences between intervention and comparison groups are indicated using asterisks (*). For each calculated difference, we also show in brackets its standard error (i.e. estimated variation in the sample), which is used in the calculation of the statistical significance level.

Column 1 shows that at endline we find an outcome X for the intervention group of NPR 150 that is significantly higher than the value of NPR 50 that is observed for the comparison group, i.e. we observe a significant difference of NPR 100. Ignoring the differences between both groups at baseline, the naïve researcher would conclude that the project has had an impact on outcome X. However, looking at the baseline values, we find that the difference of NPR 100 was already present at baseline, i.e. prior to the project intervention, even after using PSM techniques. This example illustrates the importance of looking at differences in *changes* between matched intervention and comparison groups (i.e. PSM *differences-in-differences*), which is highlighted in column 1 to be insignificant at 0. We conclude that the project had no significant impact on outcome X.

Next, column 2 shows the results for outcome Y. Here, the naïve researcher ignores the trend experienced by the comparison group and concludes that the project has had an impact on the beneficiary group of NPR 150, which represents the observed change in outcome Y for the beneficiary group. However, taking into consideration the doubling of the comparison group's outcome Y from 148 to 296 we conclude that the observed difference in changes over time, i.e. 2, is not significantly different from zero. The project had no significant impact on outcome Y.

Finally, column 3 shows the results for outcome Z. We find that the intervention group and matched comparison group were very similar at baseline, with values of respectively 150 and 148 (not significantly different). At endline, whereas the outcome of the intervention group has doubled to reach NPR 300, the outcome of the comparison group has remained fairly similar at 150. In this case, we find that the PSM difference-in-difference estimate of 148 is significantly different from zero. We conclude that the project has had a significant impact on outcome Z.

Each table also shows the robustness of the results to the use of the standard OLS difference-in-difference regression techniques – using matching weights –- controlling for a set of observed time-variant characteristics. You will find that almost all our results presented in this report are robust to this robustness check.

Finally, note that for some outcomes, especially the intermediate ones, baseline data is not available and hence we cannot compute the more rigorous difference-in-difference estimate of impact. In such incidences, the respective fields will be left blank and a caveat is added to the interpretation of the results.

Appendix A3. Descriptive Statistics on Characteristics of Other Beneficiaries

Table 21. Descriptive statistics on demographic characteristics of other beneficiaries (endline)

	Old members	New members	None-member suppliers
Average HH size	5.8	5.1***	4.9***
Average number of adults in HH	3.8	3.3***	3.4**
% households with head in productive health status	89.4	90.0	88.4
Average number of young children (< 5 years) in HH	1.2	1.1	0.9*
% female headed HHs	9.1	17.0*	22.3***
% elderly HHs	24.2	19.0	22.3
% HHs with head with secondary education	32.6	29.0	24.3
% HHs involved in agriculture	98.5	96.0	94.2*
% HHs involved in income generation activity	18.2	19.0	8.7**
% HHs with member involved in casual labour	22.0	24.0	27.2
% HHs with member employed in full time job	28.0	23.0	20.4
Average age HH head	49.8	45.7**	49.5
Average number of HH members who are member of the cooperative	1.1	1.1	0.0
# observations sample	132	100	100

Notes: HH = household; asterisks indicate significance level of differences with intervention group; ** significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.*

Table 22. Descriptive statistics on access to cooperative services by other beneficiaries (endline)

		Old members	New members	Non-member suppliers
% HHs reporting having made use of service provided by cooperative over the last 12 months	Training improved agricultural methods	62.9	64.0	49.5**
	Advice and technical support	75.8	77.0	74.8
	Provision agricultural inputs	81.1	83.0	78.6
	Provision fertilisers	81.1	70.0*	64.1**
	Provision production loans	54.6	58.0	41.8*
	Women PLCs	41.7	42.0	25.2***
% HHs reporting that service provided was very good (rather than 'good' or 'not good')	Training improved agricultural methods	27.7	32.8	45.1**
	Advice and technical support	40.2	37.1	44.2
	Provision agricultural inputs	39.3	42.2	38.3
	Provision fertilisers	32.7	38.6	50.0**
	Provision production loans	23.9	35.1	43.9
	Women PLCs	22.2	43.9**	20.0
% HHs reporting having adopted improved agricultural practice	Production commercial crops/seeds	87.1	95.0**	93.2
	Inorganic fertiliser consumption	89.4	80.0**	84.5
	Compost-making	69.7	65.0	45.6***
	Use of animal urine	52.3	48.0	30.1***
	Use of improved animal manure	41.7	28.0**	18.5***
	Irrigation	89.4	83.0	84.5
# observations sample		132	100	100

Notes: HH = household; Asterisks indicate significance level of differences with intervention group; ** significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level.*

Table 23. Descriptive statistics on intermediate outcomes of other beneficiaries (endline)

		Old members	New members	Non-member suppliers
Area of land owned (hectares)		43.6	43.4	52.4
Area of land used for farming (hectares)		47.0	45.6	55.9
% farmed land allocated to cooperative product		22.3	20.1	18.7
# male HH members involved in agriculture		1.29	1.04**	1.15
# female HH members involved in agriculture		1.64	1.37**	1.40**
On-farm men-day hires		4.03	1.63	3.21
On-farm women-day hires		4.30	2.46*	3.40
% HHs mainly depend on agriculture		75.8	73.0	72.8
% HHs member farmer group		58.3	4.0***	92.2***
% HHs reporting no change in yields		18.9	21.0	18.5
% HHs reporting an increase in yields		81.1	79.0	81.6
% HHs receiving dividends		5.3	0.0**	0.0**
% HHs reporting to have accessed loans		23.5	34.0*	13.6*
Interest rate		15.94	16.09	16.21
Savings male		221	156	53***
Savings females		268	687**	268
% HHs selling to cooperative		79.6	74.0	86.4
% HHs actual produce sold to cooperative		63.0	61.1	80.7***
% HHs selling cooperative product to cooperative		77.3	75.0	77.7
Gross value sales to cooperative (12 months) NPR		10,620	9,069	9,706
% HHs selling cooperative product to other merchants		15.2	11.0	7.8
Gross value sales to other merchants (12 months) NPR		1,032	792	514
Total values sales NPR		11,653	9,861	10,219
% HHs reporting [change] in the value of sales of cooperative product	No change	26.5	16.0*	20.4
	Increase (less than double)	63.6	79.0**	72.8
	More than double increase	9.9	5.0	6.8
% HHs reporting [change] in the value of sales of other products	No change	34.1	24.0*	24.3
	Increase (less than double)	58.3	74.0**	72.8**
	More than double increase	7.6	2.0*	2.9
# observations sample		132	100	100

Notes: HH = household; asterisks indicate significance level of differences with intervention group; ** significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.*

Table 24. Descriptive statistics on final outcomes of other beneficiaries (endline)

			Old members	New members	Non-member suppliers
Total per capita expenditures per day (NPR)			219.8	261.6***	214.4
Food per capita expenditures per day (NPR)			85.1	103.4***	88.0
% households below £1 per day			1.0	0.0	3.9
% households below $2 per day			9.1	4.0	4.9
Asset index			-0.13	0.16	0.02
% food secure households			100	99	99
Total number of meals per day		Males	4.45	4.36	4.13*
		Females	3.76	3.87	3.63
Number of main meals per day		Males	2.89	2.82	2.8
		Females	2.78	2.89	2.87
Variety of food types		Males	5.26	5.11	4.67***
		Females	4.96	4.85	4.73
Subjective perception of well-being	Breaking even or doing well		93.2	93.0	92.2
	Struggling		6.8	7.0	7.8
	Breaking even		56.8	56.0	45.6*
	Doing well		36.4	37.0	46.6
Perceptions of women's economic roles		Women	70.0	71.3**	71.2**
		Men	68.3	69.7	71.7**
Intra-household decision-making score for women			92.9	92.9	95.0*
Community influence score for women			40.7	42.6***	43.1***
Cooperative influence score for women			44.6	48.0***	48.2***
Self-efficacy score for women			18.8	19.7**	17.8**
# observations sample			132	100	100

Notes: HH = household; asterisks indicate significance level of differences with intervention group; ** significant at the 10 per cent level; ** significant at the 5 per cent level; *** significant at the 1 per cent level.*