Big Push for the Rural Economy: Final Impact Evaluation Report

Ali Cheema Asim I. Khwaja M. Farooq Naseer Jacob N. Shapiro Joshua Gill [*]

February 29th, 2020

[*] We thank Yanchuan Liu for outstanding research assistance in Harvard and Wafa Masood Khan, Umair Mazher, and Zain Asghar Abbasi for outstanding research assistance in Lahore.

Cheema: Associate Professor of Economics, Lahore University of Management Sciences & Senior Research Fellow, Institute of Development and Economic Alternatives; ali.cheema@cerp.org.pk. **Khwaja:** Professor of Public Policy, Harvard Kennedy School; khwaja@hks.harvard.edu. **Naseer:** Assistant Professor of Economics, Lahore University of Management Sciences; farooqn@lums.edu.pk. **Shapiro:** Professor of Politics and International Affairs, Princeton University; jns@princeton.edu. **Gill:** Graduate Student Fellow, CERP; joshua.gill@cerp.org.pk

Contents

Executive Summary

The *Big Push for the Rural Economy* (BPRE) scheme was designed to increase training value-added and productivity in four high-poverty districts of South Punjab. Trainings were offered in Bahawalpur, Bahawalnagar, Lodhran and Muzaffargarh districts (referred to as 'PEOP districts' in this report) to address skill gaps in the agriculture and livestock value chains. These trainings were conducted at the village level and aimed to saturate and diffuse best skills and practices throughout the village-level agriculture and livestock value chains. The BPRE scheme was implemented by the Punjab Skills Development Fund (PSDF) as part of the Skills Development Program (SDP), an extension of the Punjab Economic Opportunities Program (PEOP), jointly funded by the Government of Punjab (GoPb) and the United Kingdoms' Department for International Development (DfID).

PSDF launched the BPRE scheme in 2016, providing agriculture and livestock skills trainings to individuals in the PEOP districts. The general courses offered under agriculture focused on the most common crops grown in the region: wheat and cotton. The courses under livestock focused on large dairy animals, as majority of households own these animals. In addition, complementary specialised trainings were given to farm electricians and farm machinery mechanics in the agriculture value chain, and Artificial Insemination Technicians (AITs), farm supervisors, animal health workers, and Village Milk Collectors (VMCs) in the livestock value chain. [1] In addition, a linkage component in the form of village *melas* (fairs) was introduced to reduce the search cost between farmers and the specialised trainees providing services in the agriculture and livestock value chain.

To evaluate the BPRE scheme, PSDF entered into a collaborative agreement with the Centre for Economic Research in Pakistan (CERP) to conduct a rigorous impact evaluation. The objective of the collaboration was to improve the returns of PSDF's BPRE program through evaluation (and re-calibration, if necessary). This report presents the findings from the final impact evaluation of the BPRE scheme, which was conducted by CERP using the randomised-controlled-trial (RCT) methodology.

The evaluation sample for the BPRE scheme comprises of a representative sample of randomly selected households from 90 villages in the four districts of interest. Out of these 90 villages, 60 villages were randomly assigned to treatment group and 30 villages to control group. Treatment villages were offered training courses in agriculture and livestock-related sectors, while control villages were not offered any training. In a sub-sample of treatment villages, a linkage component in the form of village *melas* (fairs) was also implemented.

This report presents the findings of the impact of the BPRE scheme on outcomes such as total crop and milk production, yields, proportion of households engaged in crop/milk

[1] AITs assist with livestock breeding by providing artificial insemination (AI) service to farmers. Farm supervisors manage the operations of a small farm, their responsibilities include farm up-keep, nutrition and health management, etc. Animal health workers provide basic services and give husbandry advice to livestock keepers. VMCs collect milk from households and run quality checks before delivering it to Milk Collection Centres. All these specialised service providers are a part of the livestock value chain, which is shown in figure 4, section 5.1.2.

production, and total value of Agri-livestock output.[2] We evaluate the impact of the BPRE scheme immediately after the scheme ended in 2018 and one year later in 2019. Intention-to-Treat (ITT) effect was estimated to capture the impact of the BPRE scheme on an average household in the treatment villages. Results immediately after the training are summarised below. These results denote the treatment effect for an average household in treatment villages compared to the average household in control villages.

- We find an increase in the quantity produced of wheat (41%), cotton (43%) and milk (17%)
- An increase in yields for wheat (6%), cotton (13%) and milk (4.8%)
- An increase in the probability of household engagement in production for wheat (5.9%), cotton (8.3%), and milk (4.1%)
- An increase of 100.6% in the total value of Agri-livestock output (value of farm produce)
- An increase of 0.22 standard deviation in the advanced knowledge of agricultural best practices
- We find no statistically significant impact on knowledge of livestock best practices. Similarly, no impact on implementation of best practices in both agriculture and livestock.

Results one year after the training are summarised below. The results show that the impact of the BPRE scheme declined significantly over a one-year period (from 2018 to 2019) for majority of outcomes of interest, except milk yields. However, it is important to note that although the size of the impact decays after one year, the positive impact of the training persists for some outcomes of interest. Again, these results denote the treatment effect for an average household in treatment villages compared to the average household in control villages.

- We find higher quantity produced of wheat (17%), cotton (17.6%) and milk (6%) by an average household in treatment villages compared to the average household in control villages.
- Higher probability of household engagement in production for wheat (3.4%) and cotton (5.6%). No statistically significant impact for milk in one year after trainings
- Higher milk yields by 6%. No statistically significant impact on wheat and cotton yields
- An increase of 0.055 standard deviation in general livestock knowledge and an increase of 0.06 standard deviations in advanced livestock knowledge. No statistically significant impact on knowledge of agricultural best practices.
- No statistically significant impact on implementation of best practices in both agriculture and livestock.

[2] Total value of Agri-livestock output is constructed using the annual total wheat and cotton quantity produced and their respective median prices at the village level. It also includes the milk quantity produced from cows and buffaloes for each household, over the period of one year and its median price at the village level.

- A 0.063 standard deviation increase in financial satisfaction for the average male. No significant impact on financial satisfaction for females[3]
- No statistically significant impact on Total value of Agri-livestock output one year after the trainings
- We are unable detect any significant impact on availability, accessibility, or quality of specialised service providers[4]
- No statistically significant impact on the psychological well-being (K6 index) for both males and females

As a part of the BPRE scheme, a sub-sample of treatment villages, also received a linkage treatment in which village 'melas' (fairs) were conducted. The purpose of these melas was to enhance the linkage between trained farmers and other agents in the agriculture and livestock value chain. Our results show that the village melas had no significant additional impact over and above that of training. In other words, villages in which village melas were conducted gained no additional benefit in terms of production, extensive margins, yields or income, when compared to villages where only the trainings were conducted. This may be because these linkages are already reasonable enough or that they matter less.

In terms of the overall benefit-cost calculus, we find that this program is extremely favourable. The overall benefits for an average household in treatment villages in the first year (2018) is estimated to be PKR 96,322 – a return that even by the first year is close to 5.4 times the cost outlay (PKR 17,901) of the program![5] While the overall benefit in the second year (2019) drops to PKR 36,720, it is still substantial. In fact, even if we assume the yearly decay rate implied by these two numbers, we see that the projected benefit from the BPRE scheme for an average household in the treatment villages over a ten-year period is about PKR 155,600 – 8.7 times the initial cost outlay!

We also estimate the cost-benefit for wheat, cotton and livestock trainings separately. We find a total benefit of PKR 31,621, PKR 34,418 and PKR 27,601, respectively, immediately after the trainings (2018). This gives a return of 1.35 times relative to the average cost per household of the wheat trainings, 1.4 times return for cotton trainings, and 1.8 times for livestock trainings. Benefits one year after the training from wheat, cotton and livestock trainings for the average household were PKR 12,551, PKR 7,884 and PKR 11,588, respectively.

Findings from the evaluation of BPRE scheme show that such big-push style trainings can have a large and substantial effect with an extremely favourable benefit-cost ratio. This bodes well for scaling such programs up to a national level.

Our results also offer several words of caution and possible additional policy interventions. Interestingly, despite the improvement in knowledge and productivity immediately after training, we were unable to detect significant changes in practices. This

[3] We evaluate the impact of the BPRE scheme on the well-being indices (K6 index, financial satisfaction index) only for 2019 (one year after the trainings) as data on these indices was not available for 2018 (immediately after the trainings). This is also the case for the analysis on the availability, accessibility and quality of specialised service providers.

[4] Refer to section 8.8 for more information on why we are unable to detect an impact for specialised providers

[5] Cost outlay of the program is the total cost in "per household" terms (PKR 17,901 per household) of the entire BPRE scheme.

suggests that either our measures of practices were not sufficiently accurate, or that perhaps the knowledge gained through trainings did not in fact change the measured practices, but rather increased the returns to their existing practices which resulted in increased output. This requires more exploration, especially if changed practices could further enhance productivity.

Second, our results show that linking agents across the agri-livestock value chain through village melas does not seem to have additional impact over and above the impact from training. This may either be because these linkages are already reasonable enough or that they matter less. From a policy perspective, this warrants further examination to see whether we need to design stronger linkage programs or that in fact these are not needed as the market naturally creates linkages as needed.

Finally, and perhaps most importantly, the impact of the training declined significantly over a one-year period. Intriguingly, this decline did not show up in milk yields, suggesting that different types of knowledge may show different levels of persistence, likely based on how regularly this knowledge is applied. Overall, this points to the concern that trainees tend to forget what they have learnt in the trainings and cannot sustain the knowledge gained from trainings and the resulting gains in production. It also points to the fact that knowledge retention for a seasonal farm activity (wheat and cotton) is harder as compared with a livestock management activity that continues all year round. This suggests that our impacts could be even larger if the program were supplemented with low cost refresher trainings and support through SMS alerts or call centres (for evidence on ICT based support, see Cole and Fernando, 2012; Larochelle et al., 2017; Casaburi et al., 2013).[6]

[6] Introduction of low-cost information and communications technology (ICT) has been shown to deliver timely, relevant, and actionable advice to farmers that can improve agricultural knowledge and yields in India and Kenya. See cited papers for more information.

1. Introduction

The *Big Push for the Rural Economy* (BPRE) scheme was designed to increase training value-added and productivity in Punjab's four high-poverty districts - Bahawalpur, Bahawalnagar, Lodhran and Muzaffargarh (referred to as 'PEOP districts' in this report). BPRE aimed to saturate and diffuse 'frontier' skills and practices (latest skills and best practices used by progressive farmers in Pakistan) within village-level agricultural and livestock value chains. Trainings offered under BPRE can be categorised as agriculture and livestock related. The general courses under agriculture focused on the most common crops grown in the PEOP districts: wheat and cotton. The courses under livestock focused on large dairy animals (such as cows and buffaloes), as the majority of households in the PEOP districts own dairy animals. Complementary specialised trainings were also given in agriculture and livestock: in the former, trainings targeted farm electricians and farm machinery mechanics; in the latter, the focus was on Artificial Insemination Technicians (AITs), farm supervisors, health workers, and Village Milk Collectors (VMCs).

To evaluate the BPRE scheme, 90 villages from the four PEOP districts were randomly assigned to treatment and control groups. Treatment villages were offered trainings courses in agriculture- and livestock-related sectors while the control group received no trainings. In a sub-sample of treatment villages, a linkage component in the form of village *melas* (fairs) was introduced to enhance the linkages between trained farmers and other agents in the agriculture and livestock value chain.

This report is the final impact evaluation of the BPRE scheme; the key questions of interest are as follows:

- Does intensive human capital infusion in agriculture and livestock sectors through training in skills and practices have a positive impact on total crop and milk production, yields, proportion of households engaging in crop/milk production, and the total monetary value of Agri-livestock output?
- Is there a need to further link trained individuals and other agents in the agriculture and livestock value chain?
- How does the impact of training persist over time?
- Is the big push-style intervention cost-effective and sustainable?

The report is structured as follows: Section 2 provides the context of the BPRE scheme. Section 3 highlights the objectives of the BPRE scheme and its relevance to the global literature. Section 4 discusses the evaluation design. Section 5 elaborates on the trainings and implementation activities under BPRE. Section 6 discusses the survey and sample as well as the baseline characteristics. Section 7 provides information on the evaluation methodology. Section 8 presents the results on the impact of the BPRE scheme on our outcomes of interest. Section 9 provides a cost-benefit analysis of the BPRE scheme. Section 10 concludes the report with the lessons learnt and recommendations.

2. Context

The Department for International Development (DFID) and the Government of Punjab (GoPb) initiated the Punjab Economic Opportunities Programme (PEOP) in 2010 with an objective to increase and diversify income earning opportunities for the poor and vulnerable, initially in 4 of the top 10 poorest districts in Southern Punjab (Bahawalpur, Bahawalnagar, Lodhran and Muzaffargarh). The programme had two main components: (i) skills development for the poor and vulnerable and (ii) livestock and dairy development for micro-small farmers.[7] To implement the skills component of PEOP, DfID and GoPb collectively established a skills financing fund -Punjab Skills Development Fund (PSDF) - in October 2010 as a not-for profit company.

Under PEOP, PSDF has designed skill development schemes based on market research and findings derived from its monitoring and evaluation activities. It has to date trained over 160,000 poor and vulnerable people (over 35% women) in 250 demand-driven and market-relevant trades across 10 sectors. The PEOP programme reached its completion in June 2016. Based on the success of the skills component under PEOP, the Skills Development Program (SDP) was set up by GoPb in collaboration with DfID as an extension of PEOP with the objective to further catalyse inclusive growth through skills development.

When PSDF began operations in late 2010, there was limited data of the scale and quality required to design skills schemes relevant for both market and trainees' needs. Consequently, PSDF engaged with the Centre for Economic Research Pakistan (CERP) to develop a baseline for skills evaluation, and evaluate the returns on two key skills training schemes funded by PSDF.[8] The key components of the collaboration between CERP and PSDF included (i) producing rigorous evidence to enable PSDF to devise evidence-based and empirically grounded skills development interventions, and (ii) monitoring and evaluation of two skills training schemes, namely *Skills for Market -Market Linkages* (SFM-ML) and the *Big Push for the Rural Economy* (BPRE).

The evaluations conducted by CERP do not include evaluation of PSDF as an organisation or an overall evaluation of PSDF's skills schemes. Rather, CERP has been contracted to conduct third-party impact evaluations of the two, above mentioned, PSDF schemes.

[7] Initially, PEOP also focused on increasing access and returns to livestock through the Livestock and Dairy Development (L&DD) component. However, L&DD was closed down following recommendations of the program's second Annual Review.

[8] CERP undertook a number of data gathering initiatives in 2012-13, including an extensive household survey covering 32,000 households in the 4 pilot districts of PEOP (Bahawalpur, Bahawalnagar, Lodhran and Muzaffargarh). Findings derived from these activities were critical in filling information gaps for PSDF and providing a baseline for measuring the impact of PEOP.

3. Big Push for the Rural Economy (BPRE)

An important policy objective of Punjab Governments' Growth Strategy is to improve the productivity of agriculture and livestock sectors due to their strong forward and backward linkages. Moreover, these sectors are also the biggest employers in these districts, and income-generating opportunities are limited outside of these sectors. To this end, *Big Push for the Rural Economy* (BPRE) program was designed by PSDF to increase training value-added and productivity in the PEOP districts by saturating and diffusing frontier skills and practices within village-level agriculture and livestock value chains.

CERP's baseline survey report (2012) showed that PSDF's existing course offerings were under-serving the needs of people engaged in these sectors because of the paucity of cost-effective, off-the-shelf providers who could supply these skills (Cheema et al., 2012).[9] The under-provision of these skills was concerning as there was tremendous dispersion in the productivity, skill-set and practices of farmers residing in the same communities (Planning Commission, 2009; Rasul et. al., 2012). These gaps are important factors responsible for sluggish output growth rates, stagnant incomes and declining total factor productivity growth rates in these sectors (Ahmed & Gautam, 2013). The baseline report on livestock and dairy also found enormous variation in milk output per animal with the 25% most productive households having productivity levels more than double those of the least 25% productive households in the same village. It recommended the need for interventions that provide information on best practices and make available basic inputs and veterinary services (Rasul et al., 2012).

The BPRE model is based off the "big push" theory of economic growth. This theory postulates that barriers to development are pervasive and diseconomies of scale and negative externalities often frustrate the development process once it begins. Hence coordinated and comprehensive investments addressing multiple constraints are simultaneously needed to push the economy on the path of sustainable development. There is a minimum amount of resources that must be devoted to break the initial inertia, just as a certain amount of speed is required for an aircraft to be airborne (Rosenstein-Rodan, 1943).

The BPRE model aimed to achieve this by saturating the village economy with frontier skills throughout the agriculture and livestock value chains to exploit complementarities and economies of scale that arose out of it (Murphy, Shleifer, & Vishny, 1989; Kremer, 1993; Sachs, 1999; Nankhuni & Paniagua, 2013). The BPRE scheme was a direct response to the poor diffusion of skills within village-level value chains in the PEOP districts. It was designed by PSDF management in collaboration with a consortium of private sector companies, engaged in agriculture and livestock in the studied area, and approved by the PSDF Board. The consortium included Engro Foods, Engro Fertilizers, and Nestle Dairy and Rural Development Foundation (DRDF). This design exercise was informed by a detailed village-level value-chain analysis and reviewed by experts who had prior experience of designing similar programs.[10] Focus groups were held in villages that mapped all nodes in the relevant value chains; identified

[9] 11% of PSDF graduates were trained in skills related to these sectors in 2016.

[10] Agri-livestock Value Chain maps showing the key nodes in which skill training was provided (highlighted in red) is presented in figures 3 and 4, in sections 5.1.1 and 5.1.2 of this report.

nodes with significant deviations from best-practice; and mapped the agents involved in these nodes. This was followed by a validation survey in which the demand for specific skills within these value-chains was elicited. This exercise was conducted in 40% of the treatment sample in the PEOP districts. The value chain mapping and analysis was conducted by CERP in collaboration with Engro Foods and Engro Fertilizers, and the final findings and methodology was peer reviewed by independent experts.

The design exercise was followed by PSDF formulating a menu of trainings in frontier skills and practices aimed at nodes of the value chains where skills gaps persist. PSDF entered into a training and extension advisory partnership with two competitively selected, private sector Training Service Providers (TSPs), Engro Foods and Star Farms, under which a curriculum was developed. The menu of trainings was designed around the frontier skills and practices that have been successfully demonstrated and adopted by progressive and corporate farmers in similar areas. The trainings under the scheme were classified into two broad categories- agriculture and livestock. The courses offered under agriculture focused on wheat and cotton, as these were the most common crops grown in the pilot districts given the agro-climatic conditions. Similarly, the courses under livestock focused on large dairy animals as the majority of households own dairy animals. The final model had the following features: (a) it saturated training in frontier skills and practices within village-level chains, with a focus on nodes that exhibited skills-gaps; (b) it augmented in-class training with practical demonstration in the village; (c) delivery was synchronised with the production cycle; (d) training delivery was in-village to lower access costs; (d) training was complemented by extension advisory services provided by trainers placed in the treatment villages by the TSPs(e); and two village *melas* (fairs) were organised in a sub-sample of villages to enhance the linkages between the different types of trainees across the agri-livestock value chain (linkage component).

The implementation of the trainings was conducted by the two competitively selected training service providers (TSPs): Engro Foods and Star Farms. The implementation took the form of a randomised control trial (RCT) because PSDF was interested in rigorously evaluating the impact of this flagship scheme.

3.1 Literature Review

In this section, we briefly outline the theoretical literature behind the "big push" approach. We also provide a review of "big push" program have been implemented in other parts of the world. Finally, we discuss how BPRE fits in the "big push" literature.

The big push model was first put forward by P.N. Rosenstein-Rodan, which essentially proposes that underdeveloped economies require coordinated complementary investments to achieve sustainable economic growth (Rosenstein-Rodan, 1943). It argues that an uncoordinated approach to investments within an economy would be an inefficient use of resources, as underdeveloped economies face numerous challenges which need to be addressed simultaneously to take advantage of complementarities and economies of scale. Building on this concept, Murphy et. al. (1989) and Kremer (1993) found that the returns to investment in skills are low if skills fail to take advantage of economies of scale or production complementarities. Hence, investments that encompasses the whole value chain of an industry/sector are necessary to push that sector towards self-sustaining growth.

As a result of this literature, various foreign aid programs and inventions have employed the big push approach under which multiple issues such as human capital, infrastructure, financial inclusion, and asset ownership are addressed simultaneously. One of the more prominent big push programs is the Millennium Village Project (MVP), which was initiated in 2005 as a rural development project operating across 12 ultra poor villages in sub-Saharan African countries. It was implemented as a set of big push investments in agriculture, education, health, infrastructure, and public administration (Wanjala and Muradian, 2013).[11] Some big push programs, such as 'Targeting Ultra Poor Program' (TUP) in Afghanistan, target disadvantaged populations across villages and provide multi-sectoral interventions (asset transfers, monthly cash stipends, skills trainings, etc) to alleviate constraints experienced by those specific populations.[12]

Impact evaluations of these programs have revealed positive impacts. Wanjala and Muradian (2013) found that the program led to significantly higher agricultural productivity (70%).[13] Other have found that the program led to significant increases in farmer productivity and income obtained from farming (Barnet et al.,2018; Nziguheba et al.,2010).[14] Impact evaluation of the TUP program in Afghanistan also found that the program led to increased

[11] The set of interventions aimed at the agricultural sectors consisted of: promotion and subsidisation of improved fertilisers and seeds, agronomy training, installation of irrigation systems, extension services, and a market linkages component

[12] In the TUPs case, it targeted ultra-poor households from different villages in Afghanistan and provided multi-sectoral interventions (asset transfers, monthly cash stipends, skills trainings, etc) to improve income generating opportunities and alleviate poverty for those ultra-poor households only.

[13] An increase of 70% in agricultural productivity (10.1 bags per hectare more than the control group), 200% increase in production margins, 78% increase in self-consumption, and 50% increase in total income, yet no increase in cash-income.

[14] Barnet et al. (2018) evaluated MVP impact in Ghana, and found an increase of 38% in agricultural productivity, of which 78% could be explained by the increase in the usage of inputs. Nziguheba et al. (2010) evaluated the impact of MVP on crop yields in 8 sites of the project and found that yields significantly increased, and even doubled in certain sites.

consumption and a fall in poverty for the targeted households (Bedoya et al.,2019).[15] However, cost-effectiveness of these programs show that they are a poor value for money and sector-based interventions may yield better and more cost-effective results (Barnet et al., 2018). Additionally, because these programs implemented multi-sectoral interventions, it is impossible to disaggregate overall impact by intervention and sector.

The BPRE scheme takes a "big push" approach to human capital development by training actors along the entire agriculture and livestock value chains compared to similar programs which target only the farmers (Khatam et al., 2013; Rejesus et al., 2012; Siddiqui et al., 2012; Todo and Takahashi, 2011; Habib et al., 2007; Khan and Iqbal, 2005; Feder et al., 2003). Farmer Field Schools (FFS), are the most common human capital interventions which bring together a group of farmers, usually led by agriculture extension workers, to learn about skills and practices best suited for their farming systems. Trained farmers are expected to become farmer-trainers and organise field schools within their communities (Carpio and Maredia, 2011). A large number of impact evaluations have looked at the impact of FFS programs around the world, however, the overall results remain mixed. Some studies find positive effects on agricultural productivity, knowledge, and income while others find no significant impact on these outcomes. Furthermore, few studies are statistically rigorous and comprehensive (Waddington et al., 2014). Moreover, an important aspect in the sustainability and cost-effectiveness of FFS is the diffusion of knowledge from trained farmers to nonparticipants in the program. Yet, meta-analysis of several studies shows there is no evidence of this diffusion to the nonparticipant farmers within the communities (Waddington et al., 2014; Carpio and Maredia, 2011; Nankhuni and Paniagua, 2013; Feder et al., 2003; Siddiqui et al., 2012).

Ultimately, the BPRE scheme is a novel program that aims to increase training value-added and productivity by saturating and diffusing frontier skills and practices at multiple nodes within the village-level agriculture and livestock value chains. By enabling coordinated investments in human capital through trainings provided to farmers as well as other agents involved in the agriculture/livestock sectors, BPRE diffuses frontier knowledge and skills at different stages of the production cycle and exploits the complementarities and economies of scale that result from the "big push" model.[16]

[15] Bedoya et al. (2019) evaluate the impact of the TUP program and find that, on average, per capita consumption increased by 30% and there was a fall in the share of households below the poverty line, compared to the control group.

[16] Trainings are provided to farmers, as well as specialised agriculture and livestock service providers. A sub sample of villages are provided a linkage component as well. Details are discussed in section 5.

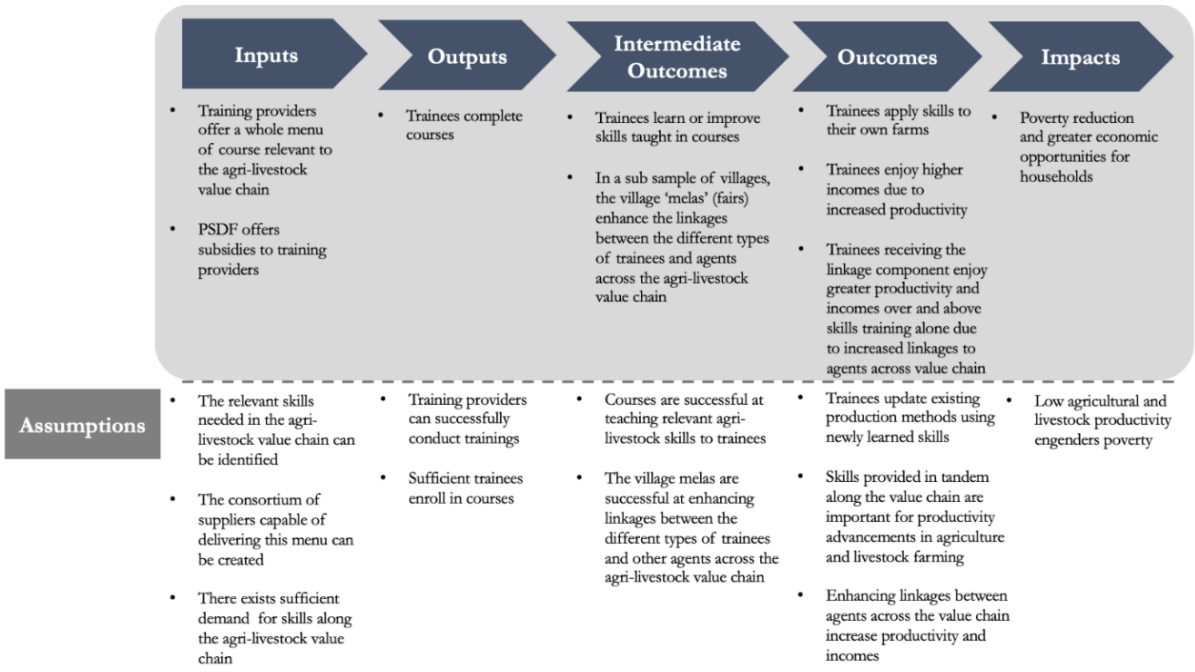
3.2 Theory of Change

As inputs to the main BPRE scheme, PSDF contracted training providers, to offer a menu of courses relevant to the agri-livestock value chains at the village level. The assumptions behind the design of this scheme were:

(i) PSDF in collaboration with the Training Service Providers (TSPs) will develop context relevant skills that the TSPs can deliver effectively.
(ii) There is sufficient demand among potential trainees for such skills.
(iii) Individuals will enrol and complete courses if offered.
(iv) Trainees will have learnt or improved upon skills taught in courses.
(v) These skills will be effective at boosting productivity.
(vi) Trainees will update existing production methods using newly learnt skills, and enjoy greater income-earning potential through productivity enhancements.
(vii) The village *melas* events (linkage component) will provide a platform for different agents in the value chain to enhance linkages which would increase productivity and income generating opportunities over and above the impact of the skills trainings alone.

The Theory of Change is also outlined in Figure 1 below.

Figure 1 Theory of Change



4. Impact Evaluation Design

The BPRE evaluation aims to measure the causal impact of "big push" style training on household productivity. In addition, it also seeks to measure the incremental impact of complementing these trainings with a linkage component. The evaluation seeks to test the empirical validity of the theory of change which proposes that human capital acquisition is particularly effective in improving productivity and catalysing growth when implemented as a "big push" across the entire value chain.

The evaluation will inform PSDF and skills development policies in Pakistan. It will also offer insights into the potential of introducing such programs in other countries as well. The results from this evaluation will shed light on the advantages of implementing a model that exploits economies of scale and production complementarities in communities. To our knowledge, this will be the first ever RCT evaluation of a novel program that aims to determine the impact of coordinated human capital investments, through the saturation of skills along the entire Agri-livestock value chain, within a village economy.

4.1 RCT Design

The BPRE scheme is being evaluated using a sample of 90 villages from the PEOP districts. Since the program is designed to operate at the community/village level, the evaluation design simply compares the average outcomes of households in villages where the program is (randomly) offered to those where the program is not.

Starting from a sample of 90 villages where CERP had conducted in-depth surveys of a large fraction of households in the village, we randomly assigned 30 villages to be control villages ("C" - where no program is offered).[17] The remaining 60 villages were assigned to two treatment groups: 30 villages were randomly assigned to just receive the menu of training ("T1" villages) and the remaining 30 villages were assigned to receive both training and trainee linkages in the form of village fairs ("T2" villages).[18] Figure 2 illustrates the evaluation design.

Since villages were randomly selected (and balance tables confirm that treatment and control villages are similar on average), we compare outcomes between the average households in T (pooling T1 and T2) and C villages, as well as between eligible and ineligible households. A household was considered eligible in agriculture (livestock) if it met at least one of the following two criteria: (i) the household grew crops (owns livestock); (ii) at least one member of the household was engaged in agriculture (livestock) related occupations. Although trainings were aimed at those already involved in the Agri-livestock sectors, ineligible households could also apply for training.

Comparing average outcomes for households in T and C villages provides an impact of the training program, while comparing T2 and T1 villages allows us to measure the incremental impact of linkages between trainees.

[17] The median village in this sample had 38% of its households surveyed

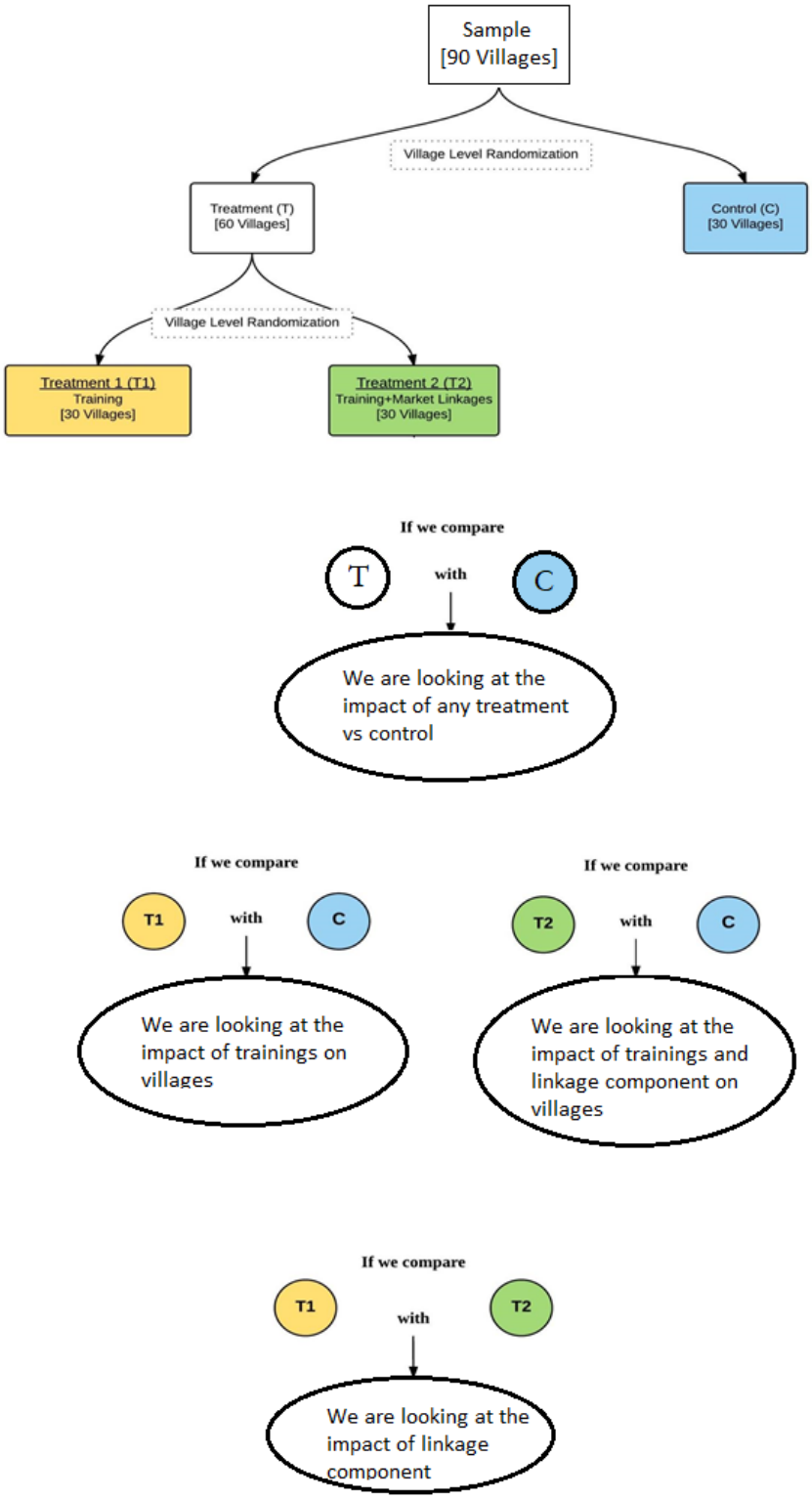
[18] Logistical and budgetary constraints did not permit including a set of villages which were only linked to the market (i.e. no training is given). Moreover, PSDF was not interested in just evaluating the impact of linkages since that is not its mandate.

4.1.1 Sample Size and Sample Selection

Within treatment villages, all households had the opportunity to enrol in courses offered by the training providers. We evaluated the impact on a sample of roughly 140 (randomly selected) households per village. Given heterogeneity in village size and logistical constraints, the final sample was adjusted down in smaller villages and up in larger villages. Therefore, we split the sample villages into four size quartiles by village population; a sample of 92 households were randomly drawn from the smallest villages, whereas villages in the largest size quartile got a sample of 188 households. The average was around 40% of the village sampled, though this naturally varied across village size. Our sample eventually amounts to around 12,700 households, with almost 70% engaged in either agriculture or livestock production.[19]

[19] We ran power calculations to determine our minimum detectable effect size using data on asset acquisition, for which we have the best data. Power calculations (with a 10 percent alpha and 80 percent power) show that we can detect an impact of more than (0.3) SD of overall training on (log) asset index. Additionally, we are powered up to detect a difference of (0.3) SD in (log) asset index between the two treatment arms.

Figure 2 Evaluation Design



5. Implementation

5.1 Details on Trainings and Linkage (Agriculture, Livestock)

The trainings offered can be classified into two broad categories: agriculture and livestock. The agriculture courses focused on wheat and cotton, as these are the most common crops grown in our pilot districts given the agro-climatic conditions. Specialised agriculture trainings were offered to specialised service providers involved in agriculture. Similarly, the courses under livestock focused on large dairy animals as majority of households own dairy animals. Specialised livestock trainings were also offered to specialised service providers involved in livestock. Table 1 shows the menu of training courses offered. Full details on the wheat, cotton, and livestock courses are provided in appendix A.

Table 1: Menu of Training Courses Offered

Courses in Agriculture	**Courses in Livestock**
1. Wheat (Seed selection/quality, land preparation, fertilizer/pesticide usage, etc.) 2. Cotton (Seed selection/quality, water management, fertilizer/pesticide usage, etc.) 3. Kitchen Gardening	1. Basic Livestock Trainings (Animal health, feed and nutrition, milking protocols and handling, breed selection, etc.)
Specialised Agriculture Trainings: 1. Farm Machinery Mechanic 2. Electrician	**Specialised Livestock Trainings:** 1. Village Milk Collection 2. Animal Health Workers 3. Artificial Insemination 4. Farm Supervisory

To ensure the skills trainings provided were effective, PSDF formulated this menu of trainings in frontier skills and practices aimed at the nodes in the value chains where skills gaps persist and have been successfully demonstrated and adopted by progressive and corporate farmers in similar areas.

Since PSDF does not conduct trainings itself, it used competitive bidding to enlist two separate training service providers (TSPs) for the agriculture and livestock components of the BPRE scheme. The bid for the agriculture component, which included the implementation of all agriculture trainings, as well as the village *melas* (fairs) for agriculture, was awarded to Star Farms.[20] The bid for the livestock component, which included the implementation of the basic livestock and specialised livestock trainings, as well as the village melas for livestock, was awarded to Engro Foods.[21]

There were two components to the trainings: a theory component, where skills were taught to trainees based on the curriculum that was developed, and a "Demonstration"

[20] Star Farms is a Metro Cash and Carry subsidiary involved in agriculture trainings in Pakistan.
[21] Engro Foods is involved in dairy products such as packaged milk. It is in a unique position of having a strong resource base with a network of trainers that can provide training in dairy.

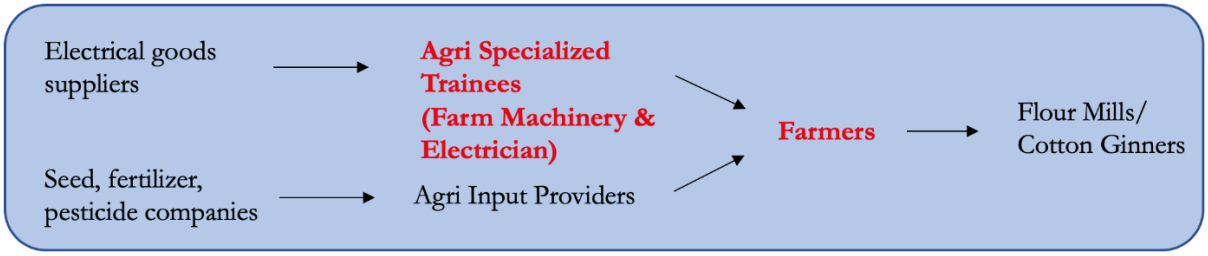
component where the skills taught in the previous component were demonstrated by the trainers to the trainees.

Additionally, half of the treatment villages (30 villages) were also offered a linkage component. This was implemented in the form of village-level *melas*, where all farmers and specialised service provider trainees were invited to a central location in the village and introduced to each other. Furthermore, downstream buyers were also invited. The objective of this exercise was to facilitate connections between trained farmers and other agents in the value chains and make them aware of the additional services available in their village that could potentially help increase their productivity, as well as linking them to potential buyers. Two village *melas* were conducted after the main training courses in agriculture and livestock had been completed: one village *mela* focused on agriculture and was conducted by Star Farms, and the other on livestock conducted by Engro Foods. The two *melas* were conducted between April and August 2018.

5.1.1 Agricultural Training

The wheat and cotton trainings were a part of the agriculture component of the BPRE scheme and were designed to be held in conjunction with the cropping cycle. To allow for flexibility, the courses were designed such that the trainees had the option of enrolling for either the cotton or wheat courses, or in cases where the farmer cultivated both crops, enrolling in both. In addition to these trainings, a kitchen gardening course was offered to women involved in small scale gardening. To infuse skills and knowledge in other nodes of the agriculture value chains, specialised service providers such as farm machinery mechanics and electricians were also offered trainings. Figure 3 shows the agriculture value chain, in which the agents that received trainings are highlighted in red.

Figure 3 Agriculture Value Chain



The wheat and cotton courses both consisted of six training modules on general agriculture, and six training modules specific to wheat or cotton. The general agriculture modules informed farmers on methods such as soil and land management techniques, farm management and record keeping, amongst others. The crop specific modules taught farmers about seed selection and quality, land preparation, and planting, amongst others.

The wheat course was designed to start with the wheat cropping cycle, which usually begins mid-November in Southern Punjab. However, due to delays in contracting the TSP and the logistics of mobilising in 60 villages, there was a delay in the start of the wheat classes, which eventually commenced on 15th December 2016. This delayed start meant that the

trainings were not synced to the first ten days of the wheat plantation cycle as originally planned, though these topics were covered during the taught course. As a remedial measure recommended by CERP, PSDF and Star Farms agreed to a follow-up with a wheat refresher in November 2017, aligned with the wheat plantation stage in the cropping cycle, to ensure that the stages in wheat plantation were practically demonstrated to the trainees.

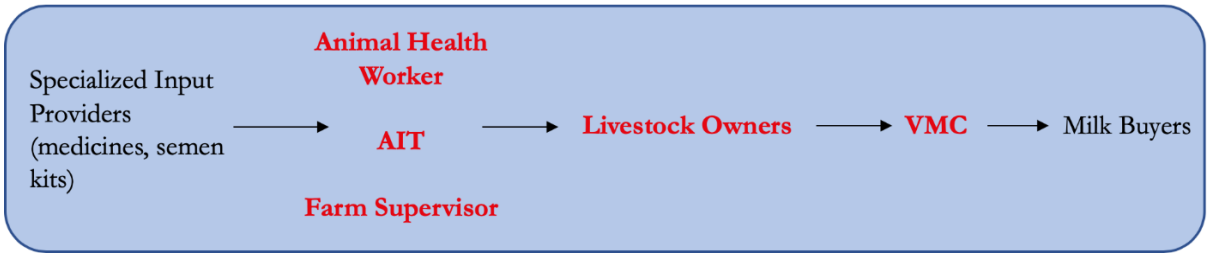
The cotton course classes began on schedule, in conjunction with the cropping cycle, during mid-May 2017, and were completed in November 2017.[22]

5.1.2 Livestock Training

The livestock component consisted of a basic livestock training course, and specialised livestock training courses for village milk collectors, animal health collectors, artificial insemination technicians, and farm technicians. The basic livestock trainings were conducted in three phases in 60 treatment villages in the four districts of Bahawalpur, Bahawalnagar, Lodhran and Muzaffargarh. The 60 treatment villages were divided into 3 phases, and trainings were conducted one phase at a time due to logistical constraints. Phase 1 was carried out in 27 villages, phase 2 in 20 villages and phase 3 in the remaining 13 villages.

The livestock courses classes were held separately for males and females. The main content of the courses was similar for both genders. Figure 4 shows the livestock value chain, in which the agents that received trainings are highlighted in red.

Figure 4 Livestock Value Chain



The basic livestock training was designed to introduce dairy farmers to animal health practices related to immunity and vaccination, deworming, and animal hygiene. There was emphasis on management of major diseases and outbreak management and stress management. The training done was both theoretical and practical in nature. A component on fodder preservation updated farmers on different fodder preservation techniques and their importance; they were also provided information on the different machinery available for using the techniques taught as part of the module.

The timelines for livestock course phases were: Phase 1: April-July 2017; Phase 2: July-September 2017; Phase 3: November 2017- March 2018.

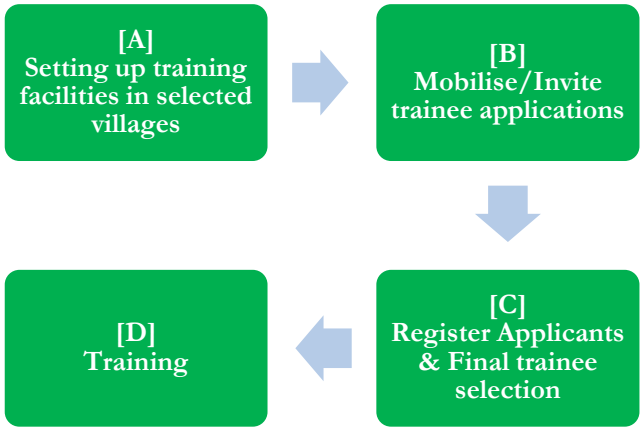
In T2 villages, Engro Foods acted as a buyer of last resort if VMCs were unable to sell all the milk they had collected.

[22] BPRE trainings rollout and completion timeline are presented in Appendix A.

5.2 Rollout Activities

Given that the BPRE scheme was subject to an RCT impact evaluation, it was important that certain protocols related to social mobilisation, enrolment and attendance verification were upheld. For this purpose, CERP audited all mobilisation activities and community sessions; oversaw applicant list generation and enrolment confirmation; maintained enrolment lists; audited class sessions; randomly surveyed class participants and collected attendance data. Also, CERP engaged with the TSPs in on-boarding sessions which provided the TSP an overview of RCT evaluations and their purpose, as well as detailing the protocols that the TSPs would be required to follow to ensure compliance with the requirements of the RCT evaluation. Furthermore, CERP outlined the specific activities that the TSP would be required to undertake during the Training Rollout (shown in Figure 5).

Figure 5: Roll Out Activities



5.2.1 Encouragement

The trainings for both agriculture and livestock targeted the eligible households in the treatment villages. A household was considered eligible in agriculture (livestock) if it met at least one of the following two criteria: (i) the household grew crops (owns livestock); (ii) at least one member of the household was engaged in agriculture (livestock) related occupations. The following strategy was used to determine the number of training slots for each village: first, based on the BPRE sample's eligible population in our baseline survey in 2016, we extrapolated the number of households in the village that were engaged in agriculture/livestock.[23] Second, a validation survey was conducted based on which we determined the level of interest for trainings. Third, we assumed that 50% of those who had shown interest would take up the trainings.[24]

To ensure that a sufficient number of eligible households in BPRE sample participated in the trainings, a voucher delivery activity was first undertaken. Vouchers were designed as

[23] Specifically, BPRE "sample and eligible households" are those who are both in our (baseline in 2016) survey sample, and are eligible for livestock/agriculture training.
[24] The assumption of a 50% take up rate was based on prior PSDF experience from other similar schemes.

an encouragement mechanism and distributed to our eligible households in the (treatment) villages and served as guarantee that these households would enjoy priority in the case of excess demand. This activity was carried out in all 60 treatment villages over the span of two weeks in October 2016. A total of 4,594 vouchers were delivered to households that were "agriculture-eligible", and 4,894 vouchers were delivered to the "livestock-eligible."

Although priority was given to voucher holders for enrolment in their desired course, any household (eligible or ineligible) could apply for trainings.[25] In practice, everyone who applied for training was accommodated.

5.2.2 Social Mobilisation

The next phase of the rollout activities was social mobilisation. Star Farms and Engro Foods, were responsible for undertaking the social mobilisation activities, which consisted of three phases; Voucher Household Visits, Community Mobilisation Sessions, and Community Awareness Activities.

During the Voucher Household Visits, the TSP Social Mobiliser ensured that each of our sample eligible households that had received vouchers for the training was encouraged to participate and invited to the community sessions.[26] The purpose of the Community Mobilisation Sessions was to provide a more detailed overview of the trainings to the villagers and facilitate them in the application process. Prospective applicants were able to fill out and submit applications or redeem vouchers. Community Awareness Activities included putting up banners and pamphlets distributed to advertise the trainings. In addition, regular announcements regarding the upcoming trainings were also made during the mobilisation period.

Applications were collected by the relevant TSPs throughout the social mobilisation phase.[27] Once the completed applications were handed over by Engro Foods/Star Farms to RCONS (survey firm hired by CERP), the RCONS enumerator then verified that all participant details were submitted and digitised the data to share with CERP.

5.2.3 Enrolment

After the collection of applications, an enrolment list for the trainings was generated to accommodate all applicants but prioritised vouchers holders, and the remaining spots were open for other applicants. Ultimately, for all the training courses, all those who had applied for the trainings were offered training. The enrolment confirmation process for the courses spanned over a week where applicants were asked to visit training centres and confirm their

[25] It was planned that in the case of excess demand, voucher holders would get priority, and the remaining slots would be allocated through a ballot. In practice, everyone was accommodated by either due to low take-up rates or by increasing number of slots where needed.

[26] The voucher households were visited by the TSP Social Mobiliser and these visits were monitored by CERP representatives. Each social mobiliser was accompanied by the CERP representative as part of this monitoring process.

[27] Applications were only accepted if they fulfilled enrollment requirements (passport size picture, a copy/picture of CNIC)

enrolments. Additionally, voucher households were also paid visits by CERP representatives asking to confirm data on enrolment.

After the enrolment confirmation, the training courses for the agriculture and livestock components began. During the training courses, CERP randomly audited class sessions, surveyed class participants, and collected attendance data in order to record voucher holder enrolment as well as drop outs in the course. Furthermore, a census activity was conducted to serve as a check on the attendance data. The purpose of the census activity was to collect basic data on trainees, identifying trainees from our BPRE sample households, regardless of whether they had been provided with vouchers.[28] Table 2 shows the BPRE sample household enrolment data for each training course.[29] Out of the 4594 vouchers delivered to BPRE sample households that were 'agriculture-eligible', 1,981 BPRE eligible sample households enrolled for wheat training, while 1,686 BPRE eligible sample households enrolled for cotton training. Hence, the uptake rate for wheat and cotton for our BPRE eligible sample households was 43% and 36% respectively. For livestock, 4,894 vouchers were delivered to BPRE eligible sample households that were 'livestock-eligible', out of which 2,717 BPRE eligible sample households enrolled in training, leading to a 56% uptake rate.

Although trainings were aimed at those already involved in the agri-livestock sectors, ineligible households could also apply for training. Table 2 also shows the take-up rates for ineligible households. Though the take up-rates were understandably low, one possible reason for ineligible households taking up the training courses may be that they are potential entrants to the livestock and agriculture sectors and require skills trainings to gain the relevant knowledge.

Table 2A: BPRE Sample Household Enrolment in Training Courses

Training Course	BPRE Eligible Sample Households	Out of Total BPRE Sample Eligible HHs (%)	BPRE Ineligible Sample Households	Out of Total BPRE Sample Ineligible HHs (%)
Wheat	1,981	43	448	12
Cotton	1,686	36	462	12
Basic Livestock	2,717	56	713	18

Note: BPRE eligible households are those that are tracked by our survey, and *either: (i) grow crops/own livestock, or (ii) have at least one member from household involved in agri-livestock sector. BPRE Ineligible households are those that are tracked by our survey, but are not involved in agri-livestock sector.*

Table 2B shows the total number of trainees under the BPRE scheme in the 60 treatment villages, including those not covered by our surveys. The total number of trainees trained under the BPRE scheme was 25,464, of which 12,778 were trained in the Wheat, Cotton and Kitchen Gardening components, while 11,514 were trained in Basic Livestock trainings. The number of trained specialised service providers was 511 and 661 for the agriculture and livestock related sectors respectively.

[28] BPRE sample households are those tracked by our survey, and include both eligible and ineligible households (those not involved in agri-livestock sectors)

[29] For more detailed overview of enrolment in training courses, refer to appendix B

In total, 26% out of the total number of trainees were female. In the agricultural trainings, females only received training in the Kitchen Gardening component. In livestock, 41% of the trainees in the Basic Livestock component were females.

Table 2B: Total Number of Trainees under BPRE Scheme by Sector

Training	Number of Trainees	Proportion of Females (%)
Agriculture (Total)	**13,289**	**14.6**
Wheat and Cotton Trainings	10,826	0
Kitchen Gardening	1,952	100
Specialised Trainings in Agriculture	511	0
Livestock (Total)	**12,175**	**38.3**
Basic Livestock Training	11,514	41.0
Specialised Trainings in Livestock	661	0
Total Number of Trainees under BPRE Scheme	**25,464**	**26**

6. Surveys and Sample

6.1 Surveys and Attrition

Household surveys were used to gather data on our outcomes of interest. Individuals in the evaluation sample were tracked through four rounds of surveys (two pre-treatment and two post-treatment rounds). We hired a local survey firm to conduct these surveys. The survey firm hired and trained their enumerators (male enumerators for male survey and female enumerators for female survey), while we monitored the trainings and field activity through spot checks. We were also provided with regular field reports during the survey activity to check for sample response rates.

As stated above, there were two pre-treatment rounds. Prior to the inception of the BPRE scheme, CERP conducted an in-depth survey of all 36,800 households in the 90 PEOP villages as a part of the PEOP program in 2013. For the BPRE scheme, this in-depth survey was updated through the 2016 baseline survey, which was conducted on a sample of the households in the 90 villages, due to logistical reasons (as discussed in section 4.2.1). Hence, the in-depth (Baseline 1) and 2016 (Baseline 2) surveys will be used as one of the two pre-treatment surveys.

After the baseline survey was conducted in 2016, a (shorter) post-treatment tracker was conducted in 2018 immediately after the BPRE trainings concluded, and an endline survey was conducted in 2019, a year after the BPRE trainings. All surveys were conducted on the full sample of 12,710 households in the four PEOP districts of Bahawalpur, Bahawalnagar, Muzaffargarh, and Lodhran.

The timeline for the four survey rounds and the BPRE trainings and linkage component was as follows:

Table 3 Timeline of Key BPRE Activities

Survey Round	Timeline
Baseline 1	2013
Baseline 2	August—October 2016
Post-treatment Tracker	February—April 2018
Endline	March—June 2019
Key Activities under the BPRE Scheme	
Agriculture Trainings	Dec 2016—Nov 2017
Livestock Training	April 2017— March 2018
Specialised Trainings	April 2017— August 2018
Village Melas (Linkage component)	April 2018— August 2018

While the 2016 baseline and post-treatment tracker surveys were conducted with the male head of the household (to minimise costs), the in-depth and endline surveys were conducted with both the male *and* female heads of the household in the BPRE sample villages. This is because we expected women to have a better understanding of some household

outcomes (such as certain livestock related practices) and men to have a good understanding of other outcomes (such as agriculture related practices).

In general, the surveys collected data on the household's employment status, income, expenditure, consumption, assets, data on milk/crop production and output levels, other data on livestock/agriculture related practices, characteristics such as land quality, access to water, and household head literacy. The female survey was shorter than the male surveys, and collected data on household (consumption, assets income), and livestock related knowledge and practices.[30]

There were a number of advantages to having a post-treatment survey immediately after the major trainings concluded, and an endline survey a year after that. Firstly, it allowed us to monitor whether any impact of the trainings on agriculture and livestock production is sustained over time. Secondly, having multiple rounds of data collection helped us to gain power to detect impact on outcomes of interest. Thirdly, by conducting the two post-treatment surveys in different seasons, we could detect the treatment impacts during high milk production (winter) and low milk production season (summer).

Post-treatment tracker survey	Endline survey
Captures immediate impact of the training on knowledge and skills	Captures whether any improvements in knowledge are retained
Captures impact of the trainings on yields and practices *during* the training period (the training was synced with the crop cycle)	Captures whether any impact on yields and practices is sustained post-training
Captures milk yields during high-yield season	Captures the impact of all specialized trainings and the market linkage scheme

The endline survey activity, originally planned to run from December 2018 to February 2019, was postponed due to delays in obtaining our No Objection Certificate (NOC) from the Government, without which it was not possible to start field activity. We were able to start the survey in three districts in early March, yet the district-level approval for Muzaffargarh took longer than the other three districts. Consequently, the endline survey only started in Muzaffargarh close to the end of March and concluded in June 2019.

Nevertheless, one issue that can arise with multiple survey rounds is sample attrition. Table 4 shows the number of households we surveyed across the different follow-ups.

[30] Females were only surveyed on Kitchen Gardening related activities for agriculture in the Endline survey. In the in-depth survey (Baseline 1), females were not surveyed on any agricultural related activities.

Table 4 Survey Attrition

Survey Round	Survey Respondents (n)	Respondents relative to Baseline 1 (%)	Respondents relative to Baseline 2 (%)	Timeline
Baseline 1	12,710	100	-	2013
Baseline 2	11,418	89.8	100	Aug—Oct 2016
Post-treatment Tracker	11,351	89.3	99.4	Feb—April 2018
Endline	11,133	87.5	97.5	March—June 2019

At the time of Baseline 2, 11,418 households out of the 12,710 sample households were surveyed. This indicates an attrition rate of 10.2%, which was mostly due to migration—6.07% of the total BPRE sample households migrated in the three years between baseline 1 and baseline 2. We also see a loss of 10.7% respondents in the post-treatment tracker, as 11,351 households were covered relative to baseline 1 (2013). However, using the households covered as part of the baseline 2 conducted in 2016 as the reference point, the coverage rate obtained for post-treatment tracker is 99.4%. In the endline survey, 11,133 households were covered, resulting in an attrition rate of 12.4% relative to the Baseline (1) survey. However, relative to the baseline 2 survey, the total attrition rate is a lot less – just 2.5% between baseline 2 and endline survey.

6.2 Baseline Characteristics

Table 5 summarises the average Baseline 2 characteristics of the households in the sample. On average, the head of households were predominately men in their late 40s with low levels of education. Only 3.9 percent of households had a female head. Around 69% of the sample households were involved in the agriculture or livestock related sectors at baseline. The average wheat yield was around 4 maund per kanal for those growing wheat, and 1.9 maund per kanal on average for those growing cotton.[31] At Baseline 2, the number of animals owned was 1.7 animals on average, with average daily milk output per animal (litres) at around 4.1 litres.

After the collection of the Baseline 2 survey in 2016, we ran balance checks on several pre-treatment outcomes between the treatment groups (T1 and T2 villages) and control group at Baseline 2 to test whether they shared similar characteristics and randomisation was successful. The sample households were balanced on the majority of pre-treatment outcomes and there was no major difference between the control and treatment groups.[32] The balance check tables are presented in Appendix C.

[31] Maund is a standardised unit used for crop produce, while Kanal is a standardised unit for land, in Pakistan. One maund equals to 40 kilograms, while 1 Kanal equals to 0.125 acre.

[32] If there was a statistically significant difference between treatment and control groups, or between T1 and T2 villages, those variables were included as controls in our regressions.

Table 5 Baseline 2016 Characteristics

	Mean	Standard Deviation	Minimum	Maximum
Household Characteristics				
HH size *(n)*	6.1	2.5	1	36
HH Head Age *(n)*	47.1	14.1	17	108
Female Head *(%)*	3.9	19.3	0	100
Involved in Agriculture* *(%)*	58.0	49	0	100
Involved in Livestock* *(%)*	51.4	49	0	100
Involved in either* *(%)*	69.3	46	0	100
Cultivable land owned or rented *(kanal)*	12.8	37.6	0	1392
Number of animals owned *(n)*	1.7	2.9	0	68
Have access to tube well *(%)*	91.5	27.8	0	100
Own a tractor (%)	2.6	15.8	0	100
Grew Fodder (%)	28	45	0	100
Land suffered salinity (%)	18.7	39	0	100
Land suffered water logging (%)	78	41	0	100
Years of experience in Wheat growing	17.28	12.31	1	90
Years of experience in Cotton growing	17.89	12.52	1	90
Production Characteristics				
Wheat yield *(maund per kanal)*	4.1	2.4	0.1	55.5
Cotton yield *(maund per kanal)*	1.9	1.3	0.01	32
Daily milk output per animal *(litres per animal)*	4.1	1.2	0	19.7

*Note: *Involvement in agriculture/livestock requires either: (i) Household grows crops/owns livestock, or (ii) has at least one member from household involved in agri-livestock related activities.*

7. Evaluation Methodology

7.1 Methods of Estimation

As the impacts of a training scheme are likely to vary for different individuals, this evaluation focuses on obtaining an average effect. Since we randomly assigned BPRE treatment at the village level, we measure the impact of the BPRE scheme in two forms: Intent-to-Treat (ITT) and Local Average Treatment Effects (LATE). An important point to note here is that both these effects do not categorically capture the effect of a particular household being trained. ITT allows for the measurement of the average effect of treatment assignment on outcomes of interest. In our case this would capture BPRE's impact on the average household in the village, regardless of whether it was involved in agri-livestock production and/or received training. In essence, this tells us the causal impact of offering, and carrying out, a range of agri-livestock trainings on the outcomes of interest.

In additional to ITT, we believe a sizable proportion of the observed impact of treatment assignment might come from people's participation in the actual training. Generally, LATE allows for evaluation of those who participated in and completed the training, and gives the average treatment effect for the treated. However, in our case, we cannot estimate the impact of any one person receiving training. Since everyone in a treatment village enjoyed access to (different types of) training around the same time, we cannot isolate the impact of any one individual being trained (from all being trained in a given village). Rather, in our case, LATE captures the contrast between no one being trained in the village to (at least one member of) *all* households in the village receiving (at least one of) the trainings. This "scaling up" of the ITT estimates is useful to give a sense of how large these effects could be under the given level of program exposure and take-up. Since LATE uses an instrumental variable (IV), we use village treatment status as an instrument for each household's participation status.

In total, we report three specifications for each outcome of interest in the results section, two of which are ITT and one is LATE. Model 1 is an ITT specification which captures BPRE's impact on the average household in the treated village, regardless of whether it was involved in agri-livestock production and/or received training.[33] Model 2 is also an ITT specification. In addition to all the variables in Model 1, Model 2 includes a dummy indicator of each household's eligibility status and an interaction between the treatment and eligibility dummies. Thereby, model 2 tells us the impact of offering, and carrying out trainings on the average eligible (and ineligible) household.[34] We run this specification as we feel they are the target population. Finally, Model 3 is a LATE specification. Recall this LATE captures the contrast between no one being trained in the village and (at least one member of) all households in the village receiving (at least one of) the trainings.

We use analysis of covariance (ANCOVA) to evaluate whether the means of an outcome variable (dependent variable) are equal (or unequal) across the treatment and control

[33] Model 1 is run on the full sample of households, controls for pre-treatment covariates, and includes a simple treatment dummy. It captures the impact, on the average household, of offering and carrying out the scheme in a treatment village.

[34] Model 2: ITT with full sample and covariates with treatment dummy interacted with dummy variable for ineligibility – the coefficient provides the ITT effect of BPRE on the average eligible household in treatment villages.

group, while controlling for the effects of other variables which vary with the dependent variable (known as covariates).[35] Standard errors are clustered at the village level for all specifications. The list of covariates used, such as socio-demographic, land characteristics, and input usage, are presented in Table D.1 in Appendix D. The results of these estimations are discussed in section 8.

7.2 Outcome Variables

Based on our theory of change, training providers disseminate skills to trainees, and the trainees learn or improve upon skills taught in courses. These skills are effective at boosting productivity as trainees update existing production methods with newly learnt skills, thereby enhancing trainees' income-earning potential.

Consequently, to investigate the impact of the trainings on productivity, we considered crop and milk production as outcome variables. Engagement in crop production, yearly crop output, and crop yield measured in maund per kanal are used as agricultural outcome variables. Engagement in crop (wheat and cotton) production is constructed as a dummy indicator, while crop output is constructed as the log of yearly crop quantity produced in maund (wheat and cotton).

Similarly, for the livestock sector, outcome variables include engagement in milk production, daily milk output levels, number of animals owned by household, and milk yield measured as milk produced per animal. Engagement in milk production is constructed as a dummy indicator, while milk production is constructed as the log of daily milk output levels in litres. Yield of milk production is measured by the daily milk output level in litres per animal. The number of animals owned is constructed by taking the log of number of dairy animals (cows and buffaloes) owned by a household. Furthermore, since females were also surveyed about livestock practices, we take the average of male and female responses for output variables, and female responses for dummy/categorical variables.[36] Log transformation applied to daily milk output levels and number of animals owned by a household.[37]

To investigate the impact of trainings on the total value of agri-livestock production, we construct the log total value of crop and milk output. It is constructed using the annual total wheat and cotton quantity produced and their respective median prices at the village level. It also includes the milk quantity produced from cows and buffaloes for each household, over the period of one year and its median price at the village level.

To examine the effectiveness of the trainings at the diffusion of knowledge and skills to the trainees, we created indices for Knowledge of Best Practices (Knowledge Index) and Input Use and Practice (Practice Index) as non-economic outcome variables of interest. For

[35] In randomised studies, ANCOVA has more power compared to ANOVA (analysis of variance) (Van Breukelen, 2006). Furthermore, McKenzie (2012) finds that ANCOVA has more power compared to the difference-in-difference (estimator).

[36] As discussed in section 6.1, we expect females to have a better understanding of livestock related activities in the household. Females were only surveyed on the livestock activities, not agriculture. Consequently, for livestock production and yield variables, we take the average of male and female responses. We primarily use female responses in dummy/categorical variables, but in the case of missing responses from females, we take the responses from males.

[37] These two measures display a log-normal distribution; hence logs are taken.

knowledge, we created two different types of indices, one for the general public and one for producers only. General Knowledge of Best Practices Indices (referred to as General Knowledge Indices) were constructed using questions that were asked of all those who were surveyed (general public); one index was created for livestock and one for agriculture. Advanced Knowledge of Best Practices Indices (referred to as Advanced Knowledge Indices) consisted of questions asked from the general public, as well as additional producer-specific questions that were asked of producers only (households involved in crop/milk production). Questions regarding Input Use and Practice (Practice Indices) were only asked of those involved in agriculture/livestock sectors, hence indices are created for producers only. We constructed each index as an additive index using multiple measures. For livestock knowledge and practice indices, we take the average of the male and female indices.[22] Table 7 shows the various Knowledge and Practice indices used, and which type of questions were used to create each.

The BPRE scheme also trained specialised technicians such as Farm Mechanics, Electricians, VMCs, Animal Health Workers, AITs, and Farm Supervisors to help facilitate production, hence we also measure the impact of the BPRE scheme on the availability, accessibility, and quality of specialised service providers. Availability of a specialised service provider is constructed as a dummy indicator, which is equal to one if at-least one service provider is working in the village, zero otherwise. Similarly, accessibility of a specialised service provider is also constructed as a dummy indicator, which is equal to one if it is easy for respondent to access the service provider, zero otherwise. Quality of service is also constructed as a dummy indicator, which equals to one if the respondent rated the quality of service to be "Very Good", and zero if the respondent rated the service as “Can be improved”. Since, the trainings of the specialised service providers were completed after the 2018 tracker survey, we are only able to evaluate the impact for 2019 (one year after the scheme ended). We use the Average Effect Size (AES) methodology which allows us to measure the mean (standardized) effect of the trainings on the availability, accessibility, and quality of the service across different service providers. We group together the specialised service providers in three categories; Agricultural specialised service providers (Farm Mechanics, Electricians), Livestock specialised service providers (VMCs, Animal Health Workers, AITs, Farm Supervisors), and All Combined (includes all types of specialised service providers).

To measure the impact that BPRE scheme may have had on the wellbeing of individuals, we constructed two wellbeing indices, the Kessler Screening Scale for Psychological Distress (k6) index and the Financial Satisfaction Index. The k6 index measures psychological distress of individuals and asks respondents how frequently they experienced the following six symptoms in the last 30 days: felt hopeless, restless or fidgety, nervous, worthless, depressed, and felt that everything was an effort. For each question, a range of 1-5 was assigned to the answer: “all of the time”, “most of the time”, “some of the time”, “a little of the time”, “none of the time”, respectively. We normalise the index so that it ranges between 0 (maximum psychological distress) and 1 (minimum psychological). The Financial-Satisfaction Index is constructed from the male and female responses on the question, “How satisfied are you with the financial situation of your household?”, on a scale between 1 and 10 with 1 being “completely dissatisfied” and 10 being “completely satisfied”. We normalise the index so that it ranges between 0 (completely dissatisfied) and 1 (completely satisfied). A point

to note is that the questions used to create both these indices were not asked in the post treatment survey (2018), hence we are only able to evaluate the longer-term impact (2019). As these questions were asked of both male and females of a household, separate indices for male and females are created.

Table 6 below presents the full list of outcome variables.

Table 6 Outcome Variables

Category	Outcome Variables
Agriculture: Crop Production	Log of yearly wheat and cotton output levels Dummy indicators of HH engagement in agricultural production Yearly wheat and cotton yields
Livestock: Milk Production	Log of daily milk output level Dummy indicators of HH engagement in milk production Daily milk output per animal Log Number of animals owned
Total Value of Agri-Livestock Production	Log of total value of agri-livestock production
General Knowledge of Best Practices *(General Public)*	Additive index of correct responses to survey questions asked to the general public, regardless of whether they are involved in crop/milk production, on knowledge about optimum frequency of Soil testing, measures to improve soil fertility, benefits of laser land levelling, how to observe land quality, measures to reduce water logging (for agriculture & livestock)
Advanced Knowledge of Best Practices *(Producers only)*	Additive index of correct responses to survey questions asked to the general public as well as on knowledge about crop patterns, correct time to plant wheat/cotton, pesticide usage, irrigation timing, fertilizer usage, calves feed/vaccination, signs of animal heat stroke, changes to animal feed in summer/winter, prevention of tick attacks, improving animal breed (for agriculture & livestock)
Input Use and Practice *(Producers only)*	Additive index based on correct responses to survey questions, from producers only, on actual practice of water use, fertilizer and pesticide use, animal feed, animal health care, milk storage, and preservation of soil quality (for agriculture & livestock)
Specialised Service Providers	Dummy indicator of service availability in the village Dummy indicator of service accessibility in the village Quality of service provided by specialised service providers We use the Average Effect Size (AES) methodology to group together specialised service providers based on three categories: Agriculture specialised service providers, Livestock specialised service providers, and Combined (Agriculture and Livestock both)
Kessler 6 Index (*Wellbeing Index*)	Kessler 6 (k6) index is an additive index which functions as a global measure of distress drawing from depressive and anxiety related symptomology. Survey respondents were asked how frequently they experienced the following six symptoms in the last 30 days: felt hopeless, restless or fidgety, nervous, worthless, depressed, and felt that everything was an effort. It is a normalised index ranging between 0 (maximum psychological distress) and 1 (minimum psychological). These questions were asked of both male and females of a household.
Financial Satisfaction Index *(Wellbeing Index)*	The Financial-Satisfaction Index is constructed from the responses "How satisfied are you with the financial situation of your household?" on a scale between 1 and 10 with 1 being "completely dissatisfied" and 10 being "completely satisfied". It is a normalised index ranging between 0 (completely dissatisfied) and 1 (completely satisfied).

Table 7 Indices Construction

Index	Sector	Questions from General Population	Questions from Producers Only
Knowledge of Best Practices		*A weighted sum of correct responses to survey questions on knowledge about:*	*Questions asked to the general public as well as questions on:*
	Agriculture *(Cotton and Wheat)*	-Concern about Climate change -Optimum frequency of Soil testing -Measures to improve soil fertility -Benefits of laser land levelling -How to observe land quality -Measures to reduce water logging -Advantages of crop rotation -Advantages of minimum tillage.	-Impact of climate change on yield -Which crops to grow -Crop patterns -Correct time to plant wheat/cotton -Pesticide usage -Irrigation timing -Fertilizer usage -Ideal seed per acre for cotton -Land preparation for rain
	Livestock	-Signs of Animal Sickness -Whether to vaccinate sick animals or not -Knowledge about Silage -Milk Chilling -Hygiene when milking animal -Cow insemination -Animal Insemination Technician -Things to consider when milking	-Signs animal has eaten enough -Daily animal water intake -Feeding milking/non-milking animals -Calves feed/vaccination -Signs of animal heat stroke -Changes to animal feed in summer -Prevention of tick attacks -Improving animal breed
Input Use and Practice		*Asked from producers only*	*Based on correct responses to survey questions on actual practice of:*
	Agriculture *(Cotton and Wheat)*	---	-Month for planting cotton/wheat -Leaving crop residue on ground -Efforts to reduce water usage -Usage of minimum tillage -Usage of organic manure -Usage of micro-nutrients -Usage of green manure -Seed per acre for wheat/cotton -Fertilizer Usage
	Livestock	---	-Whether any animals were vaccinated -Whether any animals received treatment -Who treated animal -Where animals are kept in winter/summer -Number of animals artificially inseminated -Give colostrum to new born calves -Animals access to water -How do you feed animals -How many times a day animals are fed/watered -Preservation of fodder

8. Results

In this section, we first present a summary of the results (section 8.0.1) of the impact of the BPRE scheme on our outcomes of interest, and then present detailed results (sections 8.1 onwards) using all three specification, as discussed in section 7.1. In sections 8.1 and onwards, Models 1, 2, and 3 are presented in the three columns under each outcome variable. We present both the short term (end of scheme) and longer-term impacts (one year after the scheme ended) of the BPRE scheme. Panels A, B, and C in result tables present the estimates for 2018 (short term), 2019 (one year after the scheme) and a comparison of impacts between the two years. One point to note here is that the interpretation of the coefficient for a log outcome is calculated by exponentiating the regression coefficient, subtracting one from the product and multiplying it by a hundred, thereby showing the impact in percentage.[38]

8A Summary of Results

In this subsection, we summarise the findings from the evaluation of the BPRE scheme. For simplicity, we only summarise the Intention-to-Treat estimates (Model 1) below. These results denote the treatment effect for an average household in treatment villages compared to the average household in control villages. We evaluate the impact of the BPRE scheme immediately after the scheme ended in 2018 and one year later in 2019.

Results in 2018 immediately after the training are summarised below. We find

- An increase in the quantity produced of wheat (41%), cotton (43%) and milk (17%)
- An increase in yields for wheat (6%), cotton (13%) and milk (4.8%)
- An increase in the probability of household engagement in production for wheat (5.9%), cotton (8.3%), and milk (4.1%)
- An increase of 100.6% in the total value of Agri-livestock output (value of farm produce)
- An increase of 0.22 standard deviation in the advanced knowledge of agricultural best practices
- No statistically significant impact on knowledge of livestock best practices
- No impact on implementation of best practices in both agriculture and livestock.

Results one year after the training show that the impact of the BPRE scheme declined significantly over a one-year period (from 2018 to 2019) for majority of outcomes of interest, except milk yields. However, it is important to note that although the size of the impact decays after one year, the positive impact of the training persists for some outcomes of interest. Again, these results denote the treatment effect for an average household in treatment villages compared to the average household in control villages.

[38] The formula is as follows: [*exp*(*c*) - 1] *100, where 'c' is the coefficient on the treatment variable.

- We find higher quantity produced of wheat (17%), cotton (17.6%) and milk (6%) by an average household in treatment villages compared to the average household in control villages.
- We also find higher probability of household engagement in production for wheat (3.4%) and cotton (5.6%). However, we find no statistically significant impact for milk one year after trainings.
- We find that milk yields increase by 6%; however, we see no statistically significant impact on wheat and cotton yields.
- We see an increase of 0.055 standard deviation in general livestock knowledge and an increase of 0.06 standard deviations in advanced livestock knowledge. But we find no statistically significant impact on knowledge of agricultural best practices.
- We also see a 0.063 standard deviation increase in financial satisfaction for the average male but no impact on financial satisfaction for females.[39]
- We see no statistically significant impact on Total value of Agri-livestock output one year after the trainings.
- We are unable to detect any significant impact on availability, accessibility, or quality of specialised service providers.[40]
- We also find no statistically significant impact on the psychological well-being (K6 index) for both males and females.

As a part of the BPRE scheme, a sub-sample of treatment villages also received a linkage treatment in which village 'melas' (fairs) were conducted. The purpose of these melas was to enhance the linkage between trained farmers and other agents in the agriculture and livestock value chain. Our results show that the village melas had no significant additional impact over and above that of training. In other words, villages in which village melas were conducted gained no additional benefit in terms of production, extensive margins, yields or income, when compared to villages where only the trainings were conducted. This may be because these linkages are already reasonable enough or that they matter less.

8.1 Wheat Outcomes

We measure the impact of the BPRE scheme on wheat production using three outcomes of interest: log quantity produced,[41] household involvement in wheat production (extensive margins) and yield (maund per kanal).[42]

We first present the short-term initial impact on wheat outcomes. The estimates are found in Panel A in Table 8. Recall that model 1 (column 1) is an ITT specification which tells

[39] We evaluate the impact of the BPRE scheme on the well-being indices (K6 index, financial satisfaction index) only for 2019 (one year after the trainings) as data on these indices was not available for 2018 (immediately after the trainings). This is also the case for the analysis on the availability, accessibility and quality of specialised service providers.
[40] Refer to section 8.8 for more information on why we are unable to detect an impact for specialised providers
[41] We take logs for quantity produced. Extensive margins is a dummy variable taking the value (1) if the household is involved in wheat production. Wheat yield is constructed as a ratio (quantity/land used).
[42] The take up rate for wheat training is 43% among *eligible* population and 12% among *ineligible* population.

us the causal impact of offering and carrying out the trainings on the wheat outcomes of interest for the average household. Model 2 (column 2) is also an ITT specification which tells us the causal impacts of offering and carrying out trainings on wheat outcomes of interest for the average *eligible* and *ineligible* households separately. Model 3 (column 3) is a LATE specification which captures the contrast between no one being trained in the village and (at least one member of) all households in the village receiving (at least one of) the trainings.

Generally, the results show positive impacts on all three outcomes for all three specifications, except for the impact of yield on the *ineligible* households, for which we do not find a significant impact.

In Table 8, "Ln Quantity Produced" presents the results for the impact of the BPRE scheme on wheat quantity produced. Results from model 1 (column 1) show that immediately after the completion of scheme (2018), offering trainings in the treatment villages led to an increase of 41% in wheat output for the average household. Model 2 (column 2) shows that offering trainings in the treatment villages led to an increase of 55% in wheat output for the average *eligible* household, and an increase of 24.1% for the average *ineligible* household.[43] Model 3 (column 3) shows an increase of 195% in wheat quantity produced under the hypothetical case where every household in the village was engaged in agriculture and all received training. As we noted before this is not exactly analogously to a standard LATE. In a standard LATE we would have interpreted our results as saying that the impact of training on a household *that actually received training* was production increase of 195%! The reason the LATE in our case cannot be interpreted in this way is that the treatment was randomised at the village level. Thus our LATE captures the effect of training on the trained household in the case when every household in the village (was eligible for and) also received training. Therefore, if there are complementarities in a household being trained with other households in the village, those are also being included in our estimates.

In Table 8, Panel A, "Extensive Margin" presents the results for the impact of the BPRE scheme on household involvement in wheat production. Results from model 1 (column 4) show that immediately after the completion of scheme (2018), offering trainings in the treatment villages leads to an increase of 5.9% in the probability of an average household producing wheat, when compared to the control villages. Model 2 (column 5) finds that the probability of an average *eligible* household producing wheat increases by 6.9%, by 5.1% for the average *ineligible* household. Our LATE estimate in model 3 (column 6) finds that that the probability of households producing wheat increases by 18.6%.

In Table 8, Panel A, "Yield" presents the results for the impact of the BPRE scheme on wheat yield. Our ITT model 1 (column 7) estimate shows an increase of 0.2 maund per kanal in yield for the average household, and model 2 (column 8) finds an increase of 0.3 maund per kanal in wheat yield for the average *eligible* household. No significant impact is found on the average ineligible household. Our LATE (column 8) estimate finds an increase of 0.7 maund per kanal in wheat yield under the hypothetical case whereby every household in the village was engaged in agriculture and all received training.

[43] The impact on the *ineligible* households is shown at the bottom of Panel A ("Impact on Ineligible"). The p-value below that tells us whether that estimate is significant.

The estimates of the longer-term impact (2019) are presented in Panel B in Table 8. Generally, the results show positive impacts on the wheat quantity produced and extensive margins, except for the *ineligible* households, for which we do not find significant impacts. Furthermore, we find no significant impacts on wheat yield in any of the three specifications.

In Table 8, Panel B, "Ln Quantity Produced" presents the results for the impact of the BPRE scheme on the wheat quantity produced. Results from model 1 (column 1) show that one year after the completion of scheme (2019), offering trainings in the treatment villages led to an increase of 16.9% in wheat output for the average household. Model 2 (column 2) shows that offering trainings in the treatment villages led to an increase of 25.8% in wheat output for the average *eligible* household, however, we do not find a significant impact on the ineligible households.[44] Model 3 (column 3) shows an increase of 63.2% in wheat quantity produced.

In Table 8, Panel B, "Extensive Margin" presents the results for the impact of the BPRE scheme on household involvement in wheat production. Results from model 1 (column 4) show an increase of 3.4% in the probability of an average household producing wheat, when compared to the treatment group. Model 2 (column 5) finds that the probability of an average *eligible* household producing wheat increases by 5.2%, but we do not find a significant result on ineligible households (Panel B, column 5). Our LATE estimate (column 6) finds that that the probability of households producing wheat increases by 10.7%.

In Table 8, Panel B, "Yield" presents the results for the impact of the BPRE scheme on the yield of wheat. We find not significant impact on wheat yield in the longer term.

In Table 8, Panel C shows the difference in impacts between 2018 and 2019 for all three specifications for each outcome variable.

In general, we discover significant and positive short-term impacts on all three outcomes for wheat production, and these impacts are more than twice as large as their counterparts are in 2019. Furthermore, we do not find a significant impact on wheat yield in the longer term (2019). The difference between 2018 and 2019 is statistically significant for all outcomes and specifications, except for model 2 in extensive margins (column 5, Panel C).

For the wheat quantity produced, results show a decrease in impact on an average household in treatment villages by 59% (model 1, column 1), a decrease of 53% for the eligible households in treatment villages (model 2, column 2), and a decrease of 67.7% in our LATE estimate (model 3, column 3) from 2018 to 2019.

For the impact on the likelihood of engaging in wheat production, results show a decrease of 42.4% on an average household in treatment villages (model 1, column 4), and of 42.6% for an average treatment village (model 3, column 6) from 2018 to 2019. There is no significant difference between 2018 and 2019 for eligible households in treatment villages (model 2, column 5).

The impact on wheat yield drops in 2019 by 81.1% for an average household in treatment village (model 1, column 7), 86.5% for the eligible households in treatment village

[44] The impact on the *ineligible* households is shown at the bottom of Panel B ("Impact on Ineligible"). The p-value below that tells us whether that estimate is significant.

(model 2 - column 8), and 81.5% for an average treatment village (model 3, column 9) from short (2018) to the longer term (2019).

Table 8 Results for Wheat Outcomes

	Ln of Quantity Produced			Extensive Margin			Yield (Maund per Kanal)		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE	(7) ITT	(8) ITT	(9) LATE
Panel A: Tracker 2018									
Treated	0.344***	0.438***		0.059***	0.069***		0.207***	0.318***	
	(0.084)	(0.092)		(0.019)	(0.021)		(0.054)	(0.077)	
Enrolled			1.081***			0.186***			0.650***
			(0.253)			(0.059)			(0.162)
Ineligible for Agri Training		-0.302***			-0.103***			0.104	
		(0.089)			(0.023)			(0.065)	
Treated × Ineligible		-0.222**			-0.019			-0.245***	
		(0.098)			(0.025)			(0.079)	
Baseline 2013	0.396***	0.395***	0.401***	0.359***	0.339***	0.337***	0.045***	0.046***	0.048***
	(0.037)	(0.036)	(0.038)	(0.047)	(0.047)	(0.048)	(0.012)	(0.012)	(0.012)
Baseline 2016	0.198***	0.194***	0.184***	0.175***	0.170***	0.163***	0.001**	0.001**	0.001**
	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.001)	(0.001)	(0.001)
Intercept	0.455	0.500	0.322	0.180	0.216*	0.183	2.051***	2.009***	1.934***
	(0.542)	(0.533)	(0.522)	(0.122)	(0.121)	(0.122)	(0.485)	(0.484)	(0.476)
Obs.	11351	11351	11351	11351	11351	11351	11351	11351	11351
Covariates	X	X	X	X	X	X	X	X	X
Impact on Ineligible		0.2162			0.0501			0.0723	
(p-value)		0.0275			0.0341			0.1134	
Panel B: Endline 2019									
Treated	0.156**	0.230***		0.034**	0.052***		0.039	0.043	
	(0.073)	(0.080)		(0.017)	(0.018)		(0.044)	(0.064)	
Enrolled			0.490**			0.107**			0.123
			(0.227)			(0.053)			(0.136)
Ineligible for Agri Training		-0.291***			-0.073***			0.031	
		(0.088)			(0.021)			(0.055)	
Treated × Ineligible		-0.193**			-0.042*			-0.024	
		(0.096)			(0.023)			(0.059)	
Baseline 2013	0.421***	0.420***	0.423***	0.274***	0.260***	0.260***	0.103*	0.103*	0.103*
	(0.047)	(0.046)	(0.046)	(0.052)	(0.053)	(0.051)	(0.058)	(0.058)	(0.058)
Baseline 2016	0.177***	0.173***	0.171***	0.178***	0.173***	0.171***	0.000	0.000	0.000
	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.000)	(0.000)	(0.000)
Intercept	0.433	0.469	0.350	0.164	0.186	0.162	3.138***	3.135***	3.110***
	(0.529)	(0.524)	(0.522)	(0.125)	(0.124)	(0.123)	(0.448)	(0.450)	(0.453)
Obs.	11133	11133	11133	11133	11133	11133	11133	11133	11133
Covariates	X	X	X	X	X	X	X	X	X
Impact on Ineligible		0.0367			0.0099			0.0197	
(p-value)		0.6889			0.6527			0.5633	
Panel C: Cross-Year Difference 18-19									
Diff in ITT	-0.188***			-0.025**			-0.167***		
	(0.058)			(0.013)			(0.051)		
Diff in ITT for Eligible		-0.208***			-0.017			-0.274***	
		(0.077)			(0.016)			(0.073)	
Diff in LATE			-0.591***			-0.079**			-0.528***
			(0.177)			(0.039)			(0.154)
Obs.	22484	22484	22484	22484	22484	22484	22484	22484	22484
Covariates	X	X	X	X	X	X	X	X	X
Diff in ITT Effect for Ineligible		-0.180			-0.040			-0.053	
(p-value)		0.0024			0.0059			0.2251	

Notes: Regressions of crop production variables on treatment. Standard errors clustered at the village level reported in parentheses.

8.2 Cotton Outcomes

Similar to wheat, we measure the impact of the BPRE scheme on cotton production using three outcomes of interest: log of cotton quantity produced,[45] household involvement in cotton production (extensive margins) and yield (maund per kanal).[46]

[45] We take logs for quantity produced. Extensive margin is a dummy variable taking the value (1) if the household is involved in cotton production. Cotton yield is constructed as a ratio (quantity/land used).

[46] The take up rate for cotton training is 36% among *eligible* households and 12% among *ineligible* households.

We first present the short-term impact on cotton outcomes. The estimates are found in Panel A in Table 9.

Generally, the results show positive impacts on all three cotton outcomes for all three specifications, except for the *ineligible* households, for which we do not find significant impacts on cotton yield.

In Table 9, "Ln Quantity Produced" presents the results for the impact of the BPRE scheme on cotton quantity produced. Results from model 1 (column 1) show that immediately after the completion of scheme (2018), offering trainings in the treatment villages led to an increase of 43% in cotton output for the average household. Model 2 (column 2) shows that offering trainings in the treatment villages led to an increase of 60.3% in cotton output for the average *eligible* household, and an increase of 21.5% for the average *ineligible* household. Model 3 (column 3) shows an increase of 275.8% in cotton quantity produced under the hypothetical case where every household in the village was engaged in agriculture and all received training.

In Table 9, Panel A, "Extensive Margin" presents the results for the impact of the BPRE scheme on household involvement in cotton production. Results from model 1 (column 4) show that immediately after the completion of scheme (2018), offering trainings in the treatment villages leads to an increase of 8.31% in the probability of an average household producing cotton, when compared to the control group. Model 2 (column 5) finds that the probability of an average *eligible* household producing cotton increases by 10.6%, and by 5.3% for the average *ineligible* household. Our LATE estimate in model 3 (column 6) finds that that the probability of households producing cotton increases by 30.9%.

In Table 9, Panel A, "Yield" presents the results for the impact of the BPRE scheme on the yield of cotton. Our ITT model 1(column 7) estimate shows an increase of 0.16 maund per kanal increase in cotton yield for the average household in the treatment villages. ITT model 2 (column 8) shows an increase of 0.23 maund per kanal in cotton yield for the average *eligible* household in treatment villages, whereas no significant impact is found on the average *ineligible* household. Our LATE (column 8) estimate finds an increase of 0.59 maund per kanal in cotton yield.

The estimates of the longer term impact (2019) on cotton are found in Panel B in Table 9. Generally, the results are similar to wheat outcomes, and show positive impacts on the log of cotton quantity produced and extensive margins, except for the *ineligible* households, for which we do not find significant impacts. Furthermore, we find no significant impacts on cotton yield in all three specifications.

In Table 9, Panel B, "Ln Quantity Produced" presents the results for the impact of the BPRE scheme on the cotton quantity produced in the longer term. Model 1 (column 1) show that one year after the completion of scheme (2019), offering trainings in the treatment villages led to an increase of 17.6% in cotton output for the average household. Model 2 (column 2) shows that offering trainings in the treatment villages led to an increase of 25.9% in cotton output for the average *eligible* household, however, we do not find a significant impact on the *ineligible* household. Model 3 (column 3) shows an increase of 82.6% in wheat quantity produced under the hypothetical case where every household in the village was engaged in agriculture and all received training.

In Table 9, Panel B, "Extensive Margin" presents the results for the impact of the BPRE scheme on household involvement in cotton production. Model 1 (column 4) estimates show an increase of 5.6% in the probability of an average household in a treatment village producing cotton. Model 2 (column 5) finds that the probability of an average *eligible* household in a treatment village producing cotton increases by 7.9%, but we do not find a significant result on *ineligible* households. Our LATE estimate (column 6) finds that that the probability of households producing cotton increases by 20.8%.

In Table 9, Panel B, "Yield" presents the results for the impact of the BPRE scheme on the yield of cotton. We find not significant impact on cotton yield in the longer term.

In Table 9, Panel C shows the difference in impacts between 2018 and 2019 for all three specifications for each outcome variable.

Similar to wheat outcomes, in general, we discover significant and positive short-term impacts on all three of outcomes for cotton production, and these impacts are more than twice as large as their counterparts are in 2019. Furthermore, we do not find a significant impact on cotton yield in the longer term (2019). The difference between 2018 and 2019 is statistically significant for all outcomes and specifications, except for model 2 in extensive margins (column 5, Panel c).

For the cotton quantity produced, results show a decrease in impact on an average household in treatment villages by 59% (model 1, column 1), a decrease of 57.1% for the average eligible household in treatment villages (model 2, column 2), and a decrease of 70% in our LATE estimate (model 3, column 3) from 2018 to 2019.

For the impact on the number of households producing cotton, results show a decrease of 32.7% on an average household in treatment villages (model 1, column 4), and our LATE estimates show a decrease of 32.6% (model 3, column 6) from 2018 to 2019. There is no statistically significant difference between 2018 and 2019 for the average eligible household in treatment villages (model 2, column 5).

The impact on cotton yield reduces in 2019 by 112.2% for an average household in treatment village (model 1, column 7), 116% for the eligible population in treatment village (model 2, column 8), and 112.1% in our LATE estimates (model 3, column 9) from short (2018) to the longer term (2019).

Table 9 Results for Cotton Outcomes

	Ln of Quantity Produced			Extensive Margin			Yield (Maund per Kanal)		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE	(7) ITT	(8) ITT	(9) LATE
Panel A: Tracker 2018									
Treated	0.357***	0.472***		0.083***	0.106***		0.157***	0.226***	
	(0.077)	(0.101)		(0.020)	(0.027)		(0.043)	(0.058)	
Enrolled			1.324***			0.309***			0.581***
			(0.278)			(0.073)			(0.156)
Ineligible for Agri Training		-0.136*			-0.060**			0.033	
		(0.082)			(0.026)			(0.042)	
Treated × Ineligible		-0.277***			-0.054*			-0.167***	
		(0.101)			(0.030)			(0.061)	
Baseline 2013	0.284***	0.282***	0.278***	0.175***	0.171***	0.166***	0.045**	0.045**	0.041**
	(0.039)	(0.039)	(0.040)	(0.037)	(0.036)	(0.035)	(0.018)	(0.018)	(0.018)
Baseline 2016	0.188***	0.184***	0.168***	0.178***	0.174***	0.157***	0.004***	0.004***	0.004***
	(0.015)	(0.015)	(0.016)	(0.013)	(0.013)	(0.014)	(0.001)	(0.001)	(0.001)
Intercept	0.645	0.666	0.473	0.387***	0.399***	0.352***	1.068***	1.053***	0.957***
	(0.494)	(0.489)	(0.499)	(0.114)	(0.112)	(0.113)	(0.350)	(0.352)	(0.351)
Obs.	11351	11351	11351	11351	11351	11351	11351	11351	11351
Covariates	X	X	X	X	X	X	X	X	X
Impact on Ineligible		0.1951			0.0520			0.0597	
(p-value)		0.0084			0.0146			0.1493	
Panel B: Endline 2019									
Treated	0.162***	0.230***		0.056***	0.079***		-0.019	-0.036	
	(0.057)	(0.081)		(0.014)	(0.020)		(0.034)	(0.056)	
Enrolled			0.602***			0.208***			-0.071
			(0.210)			(0.052)			(0.127)
Ineligible for Agri Training		-0.168**			-0.044**			-0.037	
		(0.074)			(0.019)			(0.049)	
Treated × Ineligible		-0.177*			-0.058**			0.040	
		(0.098)			(0.025)			(0.066)	
Baseline 2013	0.377***	0.376***	0.374***	0.147***	0.144***	0.140***	0.037*	0.037*	0.037*
	(0.041)	(0.042)	(0.041)	(0.037)	(0.037)	(0.038)	(0.021)	(0.021)	(0.020)
Baseline 2016	0.165***	0.161***	0.156***	0.168***	0.163***	0.154***	0.001	0.001	0.001
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.001)	(0.001)	(0.001)
Intercept	0.481	0.508	0.411	0.439***	0.447***	0.419***	1.393***	1.402***	1.406***
	(0.499)	(0.497)	(0.493)	(0.132)	(0.132)	(0.127)	(0.293)	(0.298)	(0.295)
Obs.	11133	11133	11133	11133	11133	11133	11133	11133	11133
Covariates	X	X	X	X	X	X	X	X	X
Impact on Ineligible		0.0530			0.0214			0.0037	
(p-value)		0.4121			0.2090			0.9101	
Panel C: Cross-Year Difference 18-19									
Diff in ITT	-0.195***			-0.027*			-0.176***		
	(0.057)			(0.015)			(0.038)		
Diff in ITT for Eligible		-0.242***			-0.027			-0.262***	
		(0.078)			(0.020)			(0.054)	
Diff in LATE			-0.722***			-0.100*			-0.653***
			(0.206)			(0.054)			(0.137)
Obs.	22484	22484	22484	22484	22484	22484	22484	22484	22484
Covariates	X	X	X	X	X	X	X	X	X
Diff in ITT Effect for Ineligible		-0.142			-0.031			-0.056	
(p-value)		0.0020			0.0170			0.0492	

Notes: Regressions of crop production variables on treatment. Standard errors clustered at the village level reported in parentheses.

8.3 Milk Outcomes

We measure the impact of the BPRE scheme on milk production through four outcomes: log of daily milk quantity produced, whether the household produces milk or not (extensive margins), milk yields (litres per animal), and log number of animals owned by the household.[47]

The results for the short-run impact on milk outcomes are found in Panel A in Table 10. In Table 10, "Ln Quantity Produced" presents the results for the impact of the BPRE

[47] The take up rate for livestock training is 56% for *eligible* population and 18% for *ineligible* population.

scheme on the milk quantity produced. Immediately after the completion of the BPRE scheme, offering trainings in the treatment villages led to an increase of 17.2% in milk quantity produced for the average household (model - column 1). Model 2 (column 2) finds an increase of 19.7% in quantity produced for the average *eligible* household, and an increase of 16.6% for the average *ineligible* household. Model 3 (column 3) shows an increase of 61.2% in milk quantity produced under the hypothetical case where every household in the village was engaged in agriculture and all received training.

In Table 10, Panel A, "Extensive Margin" presents the results for the impact of the BPRE scheme on household involvement in milk production. Results from model 1 (column 4) show that immediately after the completion of scheme (2018), offering trainings in the treatment villages leads to an increase of 4.1% in the probability of an average household producing milk, when compared to the treatment group. Model 2 (column 5) finds that the probability of an average *eligible* household producing milk increases by 3.7%, by 6.7% for the average *ineligible* household, in treatment villages. Interestingly, in the case of milk, the impact on the ineligible households is even larger than the eligible ones suggesting that some households may have been induced in producing milk when they traditionally did not. Our LATE estimate in model 3 (column 6) finds that the probability of households producing milk increases by 12.3%.

In Table 10, Panel A, "Yield" presents the results for the impact of the BPRE scheme on the yield of milk. Our ITT model (column 7) estimate shows an increase of 0.12 litres per animal increase in milk yield for the average household in treatment villages, and model 2 (column 8) finds an increase of 0.14 litres per animal in milk yield for the average *eligible* household while an increase of 0.07 litres per animal in milk yield for the average *ineligible* household. Our LATE (column 8) estimate finds an increase of 0.34 litres per animal in milk yield under the hypothetical case where every household in the village was engaged in agriculture and all received training.

In Table 11, Panel A, "Ln Number of animals" presents the results for the impact of the BPRE scheme on the average number of animals owned. Our ITT model 1 (column 1) estimate shows an increase of 7.6% increase in number of animals owned for the average household, and model 2 (column 2) finds an increase of 10% for the average *eligible* household while an increase of 7.4% in the number of animals owned for the average *ineligible* household. Our LATE (model 3, column 3) estimate finds an increase of 24.9% in the number of animals owned.

The estimates of the longer term impact (2019) on milk outcomes are found in Panel B in Table 10. We find positive impact on milk quantity produced, milk yield, and number of animals owned. The impact on the number of households involved in milk production is no longer significant in 2019 for all specifications.

In Table 10, "Ln Quantity Produced" presents the results for the impact of the BPRE scheme on the milk quantity produced. One year after the completion of the BPRE scheme, offering trainings in the treatment villages led to an increase of 6.4% in milk quantity produced for the average household in treatment villages (model 1, column 1). Model 2 (column 2) finds an increase of 10.3% in milk quantity produced for the average *eligible* household, and no significant impact on the average *ineligible* household in treatment villages. Model 3 (column

3) shows an increase of 20.2% in milk quantity produced under the hypothetical case where every household in the village was engaged in agriculture and all received training..

In Table 10, Panel B, "Extensive Margin" presents the results for the impact of the BPRE scheme on household involvement in milk production. We find no significant longer term impact on extensive margins for any of the specifications.

In Table 10, Panel B, "Yield" presents the results for the longer term impact of the BPRE scheme on the yield of milk. Our ITT model 1 (column 7) estimate shows an increase of 0.2 litres per animal increase in milk yield for the average household, and model 2 (column 8) finds an increase of 0.3 litres per animal in milk yield for the average *eligible* household, while no significant effect for the average *ineligible* household in treatment villages. Our LATE (model 3, column 8) estimate finds an increase of 0.6 litres per animal in milk yield.

In Table 11 Panel B, "Ln Number of animals" presents the results for the impact of the BPRE scheme on the average number of animals owned. We find no significant impacts in model 1 (column 1) and model 3 (column 3). Model 2 (column 2) finds an increase of 4.5% for the average *eligible* household while an increase of 7.4% in the number of animals owned for the average *ineligible* household in treatment villages.

In Table 10, Panel C shows the difference in impacts between 2018 and 2019 for all three specifications for each outcome variable. Table 11, Panel C shows the difference in number of animals owned between 2018 and 2019.

Contrary to the wheat and cotton outcomes, we find that the milk yield has increased over time. However, there is no significant impact on extensive margins, and the quantity of milk produced decreases relative to the short-term (2018). The difference between 2018 and 2019 is statistically significant for all outcomes and specifications, except for model 2 in extensive margins (column 5, Panel c) in quantity produced, extensive margins, and yield.

For the quantity of milk, model 1 (column 1) shows a decrease in impact by 63.2% for an average household in treatment villages from 2018 to 2019. Model 2 (column 2) shows a decrease of 43.8%e for the eligible population in treatment villages, while model 3 (column 3) shows a 67% decrease in milk quantity from the short-term (2018) to the longer term (2019).

For the number of households producing milk, model 1 (column 4) shows a decrease for an average household in treatment villages of 82.2%, model 3 (column 6) shows a decrease of 61.3% from 2018 to 2019. There is no statistically significant difference between 2018/19 for eligible population in treatment villages (model 2 -column 5).

For the yield of milk, model 1 (column 1) shows an increase of 83.8% for an average household in treatment villages but a decrease of 101% for the eligible population (model 2 - column 2), and a decrease of 83% in our LATE estimates (model 3, column 3) from the short-term (2018) to the longer term (2019).

The effect on number of animals owned decreases in 2019: model 1 (column 1) shows a decrease of 66.6% for an average household in treatment villages, a decrease of 54.9% for the eligible population in treatment villages (model 2, column 2), and a decrease of 68.7% in our LATE estimates (model 3, column 3) from 2018 to 2019.

Table 10 Results for Milk Outcomes

	Ln of Quantity Produced			Extensive Margin			Yield (Output per Animal)		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE	(7) ITT	(8) ITT	(9) LATE
Panel A: Tracker 2018									
Treated	0.159***	0.180***		0.041**	0.037**		0.115***	0.138***	
	(0.042)	(0.044)		(0.018)	(0.018)		(0.023)	(0.030)	
Enrolled			0.478***			0.123**			0.344***
			(0.126)			(0.054)			(0.070)
Ineligible for Lvs Training		-0.373***			-0.221***			0.091***	
		(0.045)			(0.022)			(0.032)	
Treated × Ineligible		-0.027			0.030			-0.069*	
		(0.050)			(0.023)			(0.038)	
Baseline 2013	0.331***	0.244***	0.314***	0.277***	0.190***	0.269***	0.063***	0.063***	0.062***
	(0.024)	(0.023)	(0.024)	(0.015)	(0.016)	(0.016)	(0.007)	(0.007)	(0.007)
Baseline 2016	0.000	0.000	0.000	0.186***	0.163***	0.183***	0.029***	0.029***	0.029***
	(.)	(.)	(.)	(0.012)	(0.012)	(0.012)	(0.006)	(0.006)	(0.006)
Intercept	0.198**	0.409***	0.206**	0.089**	0.218***	0.092**	2.060***	2.040***	2.041***
	(0.089)	(0.093)	(0.090)	(0.041)	(0.043)	(0.041)	(0.082)	(0.082)	(0.074)
Obs.	11351	11351	11351	11351	11351	11351	11351	11351	11351
Covariates	X	X	X	X	X	X	X	X	X
Impact on Ineligible		0.1534			0.0672			0.0694	
(p-value)		0.0015			0.0026			0.0049	
Panel B: Endline 2019									
Treated	0.062**	0.098***		0.007	0.014		0.211***	0.278***	
	(0.029)	(0.028)		(0.018)	(0.016)		(0.050)	(0.063)	
Enrolled			0.184**			0.022			0.630***
			(0.086)			(0.054)			(0.150)
Ineligible for Lvs Training		-0.334***			-0.254***			0.227***	
		(0.044)			(0.032)			(0.064)	
Treated × Ineligible		-0.053			0.013			-0.228***	
		(0.048)			(0.032)			(0.073)	
Baseline 2013	0.310***	0.227***	0.303***	0.292***	0.185***	0.291***	0.129***	0.129***	0.127***
	(0.018)	(0.017)	(0.018)	(0.013)	(0.014)	(0.013)	(0.015)	(0.015)	(0.015)
Baseline 2016	0.000	0.000	0.000	0.205***	0.177***	0.204***	0.061***	0.062***	0.062***
	(.)	(.)	(.)	(0.011)	(0.011)	(0.011)	(0.008)	(0.008)	(0.008)
Intercept	0.253***	0.443***	0.257***	0.207***	0.357***	0.207***	2.731***	2.674***	2.698***
	(0.082)	(0.082)	(0.084)	(0.041)	(0.043)	(0.040)	(0.112)	(0.116)	(0.111)
Obs.	11132	11132	11132	11132	11132	11132	11132	11132	11132
Covariates	X	X	X	X	X	X	X	X	X
Impact on Ineligible		0.0453			0.0271			0.0502	
(p-value)		0.3334			0.3903			0.3109	
Panel C: Cross-Year Difference 18-19									
Diff in ITT Effect	-0.097***			-0.034**			0.096**		
	(0.030)			(0.014)			(0.047)		
Diff in ITT Effect for Eligible		-0.082**			-0.023			0.140**	
		(0.036)			(0.015)			(0.060)	
Diff in ITT Effect for Ineligible		-0.026			-0.018			-0.159**	
		(0.039)			(0.024)			(0.061)	
Diff in LATE			-0.293***			-0.101**			0.286**
			(0.090)			(0.041)			(0.141)
Obs.	22483	22483	22483	22483	22483	22483	22483	22483	22483
Covariates	X	X	X	X	X	X	X	X	X

Notes: Regressions of milk production variables on treatment. Standard errors clustered at the village level reported in parentheses.

Table 11 Results for Number of Animals

	Number of Livestock		
	(1) ITT	(2) ITT	(3) LATE
Panel A: Tracker 2018			
Treated	0.073***	0.095***	
	(0.025)	(0.027)	
Enrolled			0.222***
			(0.076)
Ineligible for Lvs Training		-0.287***	
		(0.034)	
Treated × Ineligible		-0.024	
		(0.033)	
Baseline 2016	0.025	0.028	0.028
	(0.035)	(0.030)	(0.035)
Intercept	0.225***	0.382***	0.228***
	(0.079)	(0.078)	(0.079)
Obs.	11351	11351	11351
Covariates	X	X	X
Impact on Ineligible		0.0711	
(p-value)		0.0128	
Panel B: Endline 2019			
Treated	0.025	0.044*	
	(0.023)	(0.024)	
Enrolled			0.075
			(0.068)
Ineligible for Lvs Training		-0.238***	
		(0.034)	
Treated × Ineligible		-0.019	
		(0.034)	
Baseline 2016	0.024	0.026	0.025
	(0.033)	(0.029)	(0.033)
Intercept	0.200***	0.328***	0.201***
	(0.075)	(0.074)	(0.075)
Obs.	11132	11132	11132
Covariates	X	X	X
Impact on Ineligible		0.0248	
(p-value)		0.4179	
Panel C: Cross-Year Difference 18-19			
Diff in ITT Effect	-0.049***		
	(0.016)		
Diff in ITT Effect for Eligible		-0.051**	
		(0.019)	
Diff in LATE			-0.147***
			(0.051)
Obs.	22483	22483	22483
Covariates	X	X	X
Diff in ITT Effect for Ineligible		-0.046	
(p-value)		0.0201	

Notes: Regressions of number of livestock on treatment. Standard errors clustered at the village level reported in parentheses.

8.4 Total Value of Agri-Livestock Output

Table 12 presents a more aggregate view of the previous results by looking at the impact of the BPRE scheme on total value of agri-livestock output. It is constructed using the annual total wheat, cotton and milk quantity produced and their respective median prices at the local level. The impact on the total value of wheat/cotton/milk output separately are presented in appendix E.

The results for the short-run impact on the total value of agri-livestock output are found in Panel A in Table 12. Model 1 (column 1) results find that, immediately after the completion of the BPRE scheme, offering trainings in treatment villages leads to an increase of 100.6% in total value of agri-livestock output for the average household. Model 2 (column 2) finds an increase of 106.3% in total value of agri-livestock output for the average *eligible* household, and an increase of 146.1% for the average *ineligible* household in treatment villages. Model 3 (column 3), which presents the LATE estimates, finds an increase of 312.5% in total value of agri-livestock output under the hypothetical case where every household in the village was engaged in agriculture and received training. These numbers are quite large. However, we should note that is because the program induced a notable fraction of households who did not produce much/at all at baseline to do so (recall the extensive margin effects presented earlier). For these households, the *percentage* increase in their output is therefore extremely large, thus contributing to a large overall average percentage increase. Later on when we consider the overall benefit-cost calculation, we will show regressions at the village (rather than household level) to ensure that our overall estimated benefits accurately reflect the actual (monetary value of) benefits accruing to each household.

The results for the longer term impact on the total value of agri-livestock output are presented in Panel B, table 12. We find no significant impact in Model 1 (column 1) and Model 3 (column 3). However, model 2 (column 2) shows an increase of 34.3% in total value of agri-livestock output for the average *eligible* household, but no significant impact on the average *ineligible* household. We should note though that when we split production value for the three components (see Appendix E), even in the longer term the program impact on both wheat and cotton production value is significant (only milk production value loses significance in the longer term).

Panel C shows the difference in impacts between 2018 and 2019 for all three specifications for the total value of agri-livestock output. We find a statistically significant difference in the impact on total value of agri-livestock output between the short and longer term in all three specifications. Model 1 (column 1) shows a decrease of 79.8% for an average household in treatment villages, while Model 2 shows a decrease of 67.7% for the average *eligible* household (column 2), and a decrease of 88.1% for the average *ineligible* household in treatment villages from 2018 to 2019. Model 3 (column 3) shows a decrease of 85.4% under the hypothetical case where every household in the village was engaged in agriculture and received training.

Table 12 Results for Total Value of Agri-livestock Output

	Ln of Combined Production Worth		
	(1) ITT	(2) ITT	(3) LATE
Panel A: Tracker 2018			
Treated	0.696***	0.724***	
	(0.239)	(0.226)	
Enrolled			1.417***
			(0.477)
Ineligible for Any Training		-2.266***	
		(0.276)	
Treated × Ineligible		0.176	
		(0.327)	
Baseline 2016	0.362***	0.306***	0.350***
	(0.016)	(0.017)	(0.017)
Obs.	11351	11351	11351
Covariates	X	X	X
Impact on Ineligible		0.9007	
(p-value)		0.0143	
Panel B: Endline 2019			
Treated	0.185	0.295*	
	(0.194)	(0.151)	
Enrolled			0.376
			(0.390)
Ineligible for Any Training		-2.197***	
		(0.332)	
Treated × Ineligible		-0.135	
		(0.400)	
Baseline 2016	0.349***	0.289***	0.346***
	(0.014)	(0.015)	(0.014)
Obs.	11133	11133	11133
Covariates	X	X	X
Impact on Ineligible		0.1601	
(p-value)		0.7110	
Panel C: Cross-Year Diff 18-19			
Diff in ITT Effect	-0.511***		
	(0.177)		
Diff in ITT Effect for Eligible		-0.429**	
		(0.193)	
Diff in LATE			-1.042***
			(0.356)
Obs.	22484	22484	22484
Covariates	X	X	X
Diff in ITT Effect for Ineligible		-0.7406	
(p-value)		0.0039	

Notes: Regressions of log of combined production worth on treatment. Standard errors clustered at the village level reported in parentheses.

8.5 General Knowledge Indices

The General Knowledge questions (referred to as Knowledge of Best Practice) were asked of the general public, regardless of whether they were involved in crop/milk production. These questions were used to make General Knowledge indices. We developed two separate knowledge indices, one for livestock and one for agriculture, both of which were additive. The agriculture knowledge index was developed using questions that were asked from all households regarding soil testing, land quality, measures to reduce water logging, etc. for agriculture.[48] For livestock, the questions focused on signs of animal sickness, vaccination, milk chilling, amongst others. Because the indices are created with units that are standardised, we interpret the treatment effects as standard deviation changes in the index itself.

The general knowledge indices were only asked in 2019, so we cannot estimate the impact in the short-term. Results for the impact of BPRE scheme on general knowledge indices are present in Table 13. We find no significant impact on general agriculture knowledge indices. For general livestock knowledge indices, we find significant and positive impacts in model 1 (column 4) and model 3 (column 6). Model 1 (column 4) finds that the impact of offering and carrying out the trainings leads to an increase of 0.055 standard deviation in livestock knowledge index for the average household in treatment villages. On the other hand, model 3 (column 6) finds an increase of 0.16 standard deviation in the livestock knowledge index under the hypothetical case where every household in the village was engaged in agriculture and all received training.

[48] For a full list of areas on which the questions were asked, refer to Table 7 in section 7.2

Table 13 Results for General Knowledge Indices

	Agri Knowledge (General)			Lvs Knowledge (General)		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE
Endline 2019						
Treated	0.010	0.007		0.008*	0.007	
	(0.009)	(0.009)		(0.005)	(0.004)	
Enrolled			0.031			0.023*
			(0.025)			(0.013)
Baseline 2016	0.016	0.014	0.038**	0.003	0.003	0.009
	(0.017)	(0.017)	(0.018)	(0.010)	(0.010)	(0.010)
Intercept	0.659***	0.668***	0.657***	0.574***	0.575***	0.582***
	(0.066)	(0.067)	(0.031)	(0.015)	(0.015)	(0.015)
Obs.	11133	11133	11133	11132	11132	11132
Covariates	X	X	X	X	X	X
Impact on Ineligible		0.0114			0.0083	
(p-value)		0.3942			0.2693	

Notes: Regressions of general knowledge indices (2019 only) on treatment. Standard errors clustered at the village level reported in parentheses.

8.6 Advanced Knowledge Indices

The Advanced Knowledge questions (Knowledge of Best Practice) were asked of the producers only (households that were involved in crop/milk production). The data form these questions was used to develop Advanced Knowledge Indices. There were two separate advanced knowledge indices, one for livestock and one for agriculture. The indices were an additive index where, in addition to the questions asked of the general public, producers were asked producer-specific questions on fertiliser usage, irrigation timings, feed of animals, amongst other questions, for agriculture and livestock respectively.[49] Because the indices are created with units that are standardised, we discuss the treatment effects as units of standard deviation changes in the index itself.[50]

Model 2 is not estimated for Advanced Knowledge indices as it separates the impact on the average *eligible* and *ineligible* households. However, since these indices are developed using questions that were asked of those involved in production of crops and milk only (eligible households only), we are able to sufficiently capture the average *eligible* household impact using model 1. The results are presented in Table 14.

Table 14, Panel A presents the results for the advanced agriculture and livestock knowledge indices in the short-run. We only find a significant impact on the advanced agriculture knowledge index in the short-run only.

[49] For a full list of areas on which the questions were asked, refer to Table 7 in section 7.2

[50] To calculate the impact in standard deviations, we take the relevant coefficient and divide it by the baseline standard deviation of the respective index.

Model 1 (column 1) finds that immediately after the completion of the BPRE scheme, the impact of offering, and carrying out the trainings leads to an increase of 0.22 standard deviation in the advanced agriculture knowledge index for an average *eligible* household. Note that since these questions were only asked of eligible households Model 1 and 2 are essentially equivalent in this case and hence we do not separately run Model 2 (i.e. the sample in models 1 and 2 includes only eligible households). Our LATE (column 2) estimate finds an increase of 0.46 standard deviations in the advanced agriculture knowledge index.

Table 14, Panel B presents the longer term results for the advanced agriculture and livestock knowledge indices. There is a significant impact on the advance livestock knowledge index only in the longer term.

Model 1 (column 3) finds an increase of 0.06 standard deviations in the advanced livestock knowledge index for an average *eligible* household, while our LATE estimate (column 4) finds an increase of 0.16 standard deviations.

We find a positive impact on advanced agriculture knowledge in the short-run, and no significant impact in the longer term. Vice versa, we find a positive impact on advanced agriculture knowledge in the longer term, and no significant in the short-run.

However, the difference between short and longer term impacts are only statistically significant for the advanced agriculture knowledge index, as shown in Panel C.

Table 14 Results for Advanced Knowledge Indices

	Agri Knowledge		Lvs Knowledge	
	(1) ITT	(2) LATE	(3) ITT	(4) LATE
Panel A: Tracker 2018				
Treated	0.037***		0.006	
	(0.011)		(0.007)	
Enrolled		0.078***		0.016
		(0.023)		(0.016)
Baseline 2016	0.047***	0.049***	0.002	0.003
	(0.013)	(0.013)	(0.009)	(0.009)
Intercept	0.542***	0.532***	0.585***	0.585***
	(0.050)	(0.054)	(0.017)	(0.017)
Obs.	8118	8118	7801	7801
Covariates	X	X	X	X
Panel B: Endline 2019				
Treated	-0.001		0.009**	
	(0.007)		(0.004)	
Enrolled		-0.002		0.023**
		(0.015)		(0.010)
Baseline 2016	0.011	0.011	0.008	0.008
	(0.013)	(0.013)	(0.007)	(0.007)
Intercept	0.478***	0.479***	0.560***	0.560***
	(0.051)	(0.051)	(0.011)	(0.010)
Obs.	7070	7070	8685	8685
Covariates	X	X	X	X
Panel C: Cross-Year Difference 18-19				
Diff in ITT Effect 18-19	-0.038***		0.002	
	(0.010)		(0.006)	
Diff in LATE 18-19		-0.081***		0.007
		(0.021)		(0.016)
Obs.	15188	15188	16486	16486
Covariates	X	X	X	X

Notes: Regressions of knowledge indices (producers only) on treatment. Standard errors clustered at the village level reported in parentheses.

8.7 Practice Indices

The Input Usage and Practice Indices (referred to as Practice Indices) were created for livestock and agriculture producers only. The Practice Indices were created as additive index based on correct responses to survey questions on actual practice of water use, fertiliser and pesticide use, animal feed, animal health care, milk storage, and preservation of soil quality (for agriculture & livestock). Because the indices are created with units that are standardised, we discuss the treatment effects as units of standard deviation changes in the index itself.

The results are presented in Table 15. No significant impact on practice shows up in the longer term or short-term. The lack of a significant impact on practices is somewhat surprising since we see increases in production, extensive margins, and yield for all outcomes in either the short or longer term. This may be perhaps due to the fact that our practice measures do not capture all changes that could have occurred or there is a lot of noise in recording these practices.

Table 15 Results for Practice Indices

	Agri Practice		Lvs Practice	
	(1) ITT	(2) LATE	(3) ITT	(4) LATE
Panel A: Tracker 2018				
Treated	0.003		0.008	
	(0.010)		(0.011)	
Enrolled		0.005		0.019
		(0.019)		(0.026)
Baseline 2016	0.029**	0.029**	0.044***	0.044***
	(0.014)	(0.014)	(0.012)	(0.012)
Intercept	0.360***	0.360***	0.385***	0.386***
	(0.063)	(0.063)	(0.024)	(0.024)
Obs.	6121	6121	7140	7140
Covariates	X	X	X	X
Panel B: Endline 2019				
Treated	-0.002		-0.006	
	(0.011)		(0.011)	
Enrolled		-0.005		-0.015
		(0.023)		(0.027)
Baseline 2016	0.016	0.016	0.059***	0.060***
	(0.018)	(0.018)	(0.016)	(0.016)
Intercept	0.271***	0.272***	0.552***	0.552***
	(0.054)	(0.053)	(0.031)	(0.031)
Obs.	6781	6781	8134	8134
Covariates	X	X	X	X
Panel C: Cross-Year Difference 18-19				
Diff in ITT Effect 18-19	-0.005		-0.014	
	(0.011)		(0.015)	
Diff in LATE 18-19		-0.011		-0.034
		(0.023)		(0.039)
Obs.	12902	12902	15274	15274
Covariates	X	X	X	X

Notes: Regressions of practice indices (producers only) on treatment. Standard errors clustered at the village level reported in parentheses.

8.8 Specialised Service Providers

As mentioned earlier we also conducted trainings for specialised service providers to exploit the complementarities that existed within the value chains. However, our intervention design does not allow us to measure the direct effect of these trainings for these specialised service providers. There were several reasons why we opted for such a design. First, we maintain that the relevant and important outcome of interest is the indirect benefit of these trainings on the productivity of farmers. Second, estimating this affect would have required defining a comparison group in the control villages. This would have entailed conducting extensive additional surveys to identify similar individuals who would have potentially taken the trainings had it been offered in these village. Third, since we were training a very small number of individuals in each village (average of 3 per village), detecting an impact would have required a very large sample size, making the exercise very cost ineffective.

However, we use the following three measures to estimate the impact of training specialised service providers: Availability, Accessibility, and Quality of service. We use the Average Effect Size (AES) methodology for this and group specialised service providers using three categories: Agriculture specialised service providers (Farm Mechanics, Electricians), Livestock specialised service providers (VMCs, Animal Health Workers, AITs, Farm Supervisors), and Combined (includes all types of specialised service providers). Overall, the results are inconclusive and presented in Appendix F.

8.9 Wellbeing Indices (Kessler 6 and Financial Satisfaction)

To measure any impact that BPRE scheme may have had on the wellbeing of individuals, we constructed two wellbeing indices, the Kessler Screening Scale for Psychological Distress (k6) index and the Financial Satisfaction Index. The k6 index measures psychological distress of individuals, and asks respondents how frequently they experienced the following six symptoms in the last 30 days: felt hopeless, restless or fidgety, nervous, worthless, depressed, and felt that everything was an effort. It is a normalised index ranging between 0 (maximum psychological distress) and 1 (minimum psychological). The Financial-Satisfaction Index is constructed from the male and female responses on the question, "How satisfied are you with the financial situation of your household?", on a scale between 1 and 10 with 1 being "completely dissatisfied" and 10 being "completely satisfied". We normalise the index so that it ranges between 0 (completely dissatisfied) and 1 (completely satisfied). Both indices are created for males and females, hence we evaluate the impact separately for males and females.

We only evaluate the longer-term impact for the wellbeing indices as the questions used to construct the indices were not asked in the post treatment tracker (2018). Furthermore, because the indices are created with units that are standardised, we discuss the treatment effects as units of standard deviation changes in the index itself.

Table 18, Panel A presents the results for the male K6 index We find no statistically significant impact on the male K6 index in any of the models.

Table 18, Panel B presents the results for the female K6 index. We find no statistically significant impact on the female K6 index in any of the models.

Table 19, Panel A presents the results for the male financial satisfaction index. Model 1 (column 1) results find that one year after the completion of the BPRE scheme, offering trainings in treatment villages leads to an increase of 0.063 standard deviations in financial satisfaction for the average male in treatment villages, compared to the average male in control villages. Model 2 (column 2) find an increase of 0.065 standard deviations in financial satisfaction for the average *eligible* males in treatment villages, but no significant impact for the average *ineligible* male in treatment villages, compared to the control group. Model 3 (column 3) finds an increase of 0.121 standard deviation in financial satisfaction.

Table 19, Panel B presents the results for the female financial satisfaction index. We find no statistically significant impact on the female financial satisfaction in any of the models.

Table 16: Kessler 6 Index

	Kessler 6 Index		
	(1) ITT	(2) ITT	(3) LATE
Panel A: Male			
Treated	-0.015	-0.014	
	(0.009)	(0.009)	
Enrolled			-0.029
			(0.018)
Ineligible for Training		-0.003	
		(0.011)	
Treated x Ineligible		-0.007	
		(0.012)	
Baseline 2013	0.076***	0.076***	0.075***
	(0.015)	(0.015)	(0.015)
Baseline 2016	0.121***	0.121***	0.119***
	(0.016)	(0.016)	(0.016)
Intercept	0.502***	0.505***	0.502***
	(0.040)	(0.039)	(0.039)
Obs.	9,765	9,765	9,765
Covariates	X	X	X
Impact on Ineligible		-0.022	
(p-value)		(0.132)	
Panel B: Female			
Treated	0.001	0.003	
	(0.008)	(0.009)	
Enrolled			0.002
			(0.017)
Ineligible for Training		0.011	
		(0.009)	
Treated x Ineligible		-0.007	
		(0.011)	
Baseline 2013	0.084***	0.084***	0.084***
	(0.015)	(0.015)	(0.015)
Baseline 2016	0.118***	0.116***	0.118***
	(0.027)	(0.028)	(0.027)
Intercept	0.391***	0.389***	0.392***
	(0.036)	(0.036)	(0.035)
Obs.	10,863	10,863	10,863
Covariates	X	X	X
Impact on Ineligible		-0.004	
(p-value)		(0.777)	
Panel C: Male-Female Difference			
Diff in ITT Effect	-0.016		
	(0.013)		
Diff in Effect for eligible		-0.017	
		(0.013)	
Diff in LATE			-0.032
			(0.026)
Obs.	20,628	20,628	20,628
Covariates	X	X	X
Impact on Ineligible		-0.018	
(p-value)		(0.383)	

Table 17: Financial Satisfaction Index

	Financial Satisfaction Index		
	(1) ITT	(2) ITT	(3) LATE
Panel A: Male			
Treated	0.025**	0.026**	
	(0.010)	(0.010)	
Enrolled			0.048**
			(0.020)
Ineligible for Training		-0.000	
		(0.011)	
Treated x Ineligible		-0.008	
		(0.013)	
Baseline 2013	0.000	0.000	0.000
	(0.017)	(0.002)	(0.002)
Baseline 2016	0.060***	0.061***	0.060***
	(0.009)	(0.009)	(0.009)
Intercept	0.271***	0.272***	0.274***
	(0.036)	(0.035)	(0.034)
Obs.	9,901	9,901	9,901
Covariates	X	X	X
Impact on Ineligible		0.018	
(p-value)		(0.298)	
Panel B: Female			
Treated	0.015	0.015	
	(0.011)	(0.011)	
Enrolled			0.030
			(0.023)
Ineligible for Training		0.006	
		(0.009)	
Treated x Ineligible		0.000	
		(0.010)	
Baseline 2013	0.010***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)
Baseline 2016	0.097***	0.097***	0.096***
	(0.022)	(0.022)	(0.022)
Intercept	0.358***	0.355***	0.362***
	(0.037)	(0.037)	(0.036)
Obs.	10,939	10,939	10,939
Covariates	X	X	X
Impact on Ineligible		0.015	
(p-value)		(0.325)	
Panel C: Male-Female Difference			
Diff in ITT Effect	-0.087***		
	(0.025)		
Diff in Effect for eligible		-0.083***	
		(0.026)	
Diff in LATE			0.018
			(0.017)
Obs.	20,840	20,840	20,628
Covariates	X	X	X
Impact on Ineligible		0.003	
(p-value)		(0.785)	

8.10 Difference between T2 and T1 villages

In a subsample of villages (T2 villages), a linkage component was delivered in addition to the skills trainings, with the aim of connecting trained farmers to other agents in the agriculture and livestock value chain, such as specialised service providers and potential buyers. Based on our theory of change, if the two village 'melas' events (linkage component) is conducted to provide a platform for different agents in the value chain to enhance linkages were successful, trainees would enjoy greater productivity and income generating opportunities over and above the impact of the skills trainings alone. However, our results show that there is no statistically significant difference between T2 and T1 villages for most of the outcomes of interest, except for the livestock practice index.[51] This suggests that the melas had no significant additional impact over and above that of training.

The T2-T1 difference tables for other outcomes are presented in Appendix G.

[51] In appendix G, Table 6, "T2-T1" and "p-value" show the difference and significance of the coefficient for livestock practice index respectively. It shows a decrease in the livestock practice index when comparing T2 villages to T1 villages.

9. Cost-Benefit Analysis

PSDF used competitive bidding to contract two separate training services providers (TSPs) who were responsible for the opening and running of training centres, conducting social mobilisation, and administering of all the trainings and village *melas* (fairs).

In this section, we provide the cost-benefit analysis of the BPRE scheme. Table 20 reports the cost structure for the BPRE scheme. PSDF provided the Per Trainee Cost for the different types of trainings in agriculture and livestock, as well as the number of trainees for each type of training. The average cost of the BPRE scheme comes to PKR 17,901 ($162)[52] per household.[53]

For the big push interventions mentioned in the literature review, cost per trainee ranges from $395 (Barnett et al.,2018) to $1,614 (Bedoya et al., 2019).[54] Therefore, the cost per trainee of BPRE is significantly lower than other big push style interventions.

[52] Converted as (1$=110.63 PKR) which was the average exchange rate in 2017, the year the majority of trainings were conducted.
[53] We divide the total cost (PKR 403,256,033) by the total number of trainees (25,906) and multiply it with average number of trainees per household (1.15) to get cost at household level.
[54] In 2012 PV terms

Table 18 Cost Structure for BPRE scheme

Training	Per Trainee Cost (PKR)	Number of Trainees	Costs (PKR)
Agriculture			
Wheat & General Agriculture Training	13,321	2,787	371,25,627
Cotton & General Agriculture Training	14,022	1,270	17,807,940
Wheat, Cotton & General Agriculture Training	22,640	6,769	153,250,160
Kitchen Gardening & Farm Food Processing	5,843	1,952	11,405,536
Agriculture Extension Agent (Refresher)	9,722	222	2,158,284
Dealer (Info Session)	12,635	220	2,779,700
Farm Machinery Mechanic	35,365	360	12,731,400
Electrician	35,365	151	5,340,115
Livestock			
Basic Livestock Training	9,702	11,514	111,708,828
Extension worker Training	64,950	210	13,639,500
Farm Supervisor Training	88,143	155	13,662,165
Village Milk Collector	69,400	209	14,504,600
Artificial Insemination	82,094	87	7,142,178
Total		25,906	**403,256,033**
Average Cost per Household (PKR) 17,901			

In terms of benefits, we now run analysis at the village level for the total imputed (monetary) value of wheat, cotton and milk output generated due to the program's impact. Table 21 below is therefore the analogous to Table 12, but here we collapse everything to the village level and use level (rather than logs) of value so as to obtain an accurate estimate of the actual financial gain experienced by the average village household. We present the financial gains for each of the three components (wheat, cotton, and milk) and then put them all together in column 4 to get the combined financial gain (per household).[55] The results show that for the average household in the treatment villages, the value of annual output from agri-livestock production increases by PKR 96,322 in the first year (2018), and by PKR 36,719 in the second year (2019) (over the average household in the control villages).

[55] Recall in our discussion of Table 12, we had cautioned against using the implied percentage increases in production value at the household level to impute overall program benefits. This is because the program induced a sizable fraction of household to move from little/no production to positive production and such households understandably experienced a very large percentage increase in production (value) as they start from a very low base. This in turn leads to a large average percentage increase. While this indeed accurately reflects the *average* percentage increase across all households, imputing it on the baseline average production could lead to over-estimating actual benefits. A more conservative and accurate estimate is to instead construct the total value of production at the village level and then average this over the number of households (surveyed) in the village and run the analysis at the village level. This is precisely what we do in Table 21.

Table 19 Value of Annual Agri-Livestock Output for the average household

	Level of Income (Village Mean)			
	(1) Wheat	(2) Cotton	(3) Milk	(4) Combined
Panel A: Tracker 2018				
Treated	31621.149***	34418.466***	27601.452***	96322.415***
	(6921.986)	(10615.045)	(5244.021)	(20222.139)
Baseline 2016	0.817***	0.984***	1.128***	1.038***
	(0.112)	(0.149)	(0.125)	(0.141)
Obs.	90	90	90	90
T over Control Mean	0.5195	0.5419	0.4925	0.5339
Panel B: Endline 2019				
Treated	12551.322**	7883.926	11588.237***	36719.587**
	(5080.298)	(9687.905)	(2949.853)	(15043.914)
Baseline 2016	0.895***	1.362***	0.656***	1.125***
	(0.083)	(0.136)	(0.070)	(0.105)
Obs.	90	90	90	90
T over Control Mean	0.2062	0.1241	0.2068	0.2035
Panel C: Cross-Year Difference 18-19				
Diff in Effect	-19069.827**	-26534.540*	-16013.215***	-59602.828**
	(8586.228)	(14371.315)	(6016.759)	(25204.251)
Obs.	180	180	180	180

Notes: Regressions of village-average production (level) variables on treatment.

Considering the benefits from the first year (2018) for average households in treatment villages, we obtain a total benefit of PKR 96,322 – or a close to 5.4 times return relative to the cost outlay of the program! This is extremely favourable and does not even include the additional benefit in the second year (2019) of PKR 36,720.

We can also estimate the projected benefits from the BPRE scheme over the next ten years assuming a constant decay rate of 61.88% per annum.[56] Using this we estimated projected benefit from the BPRE scheme for an average household in the treatment villages is PKR 155,600 in terms of increased value of annual agri-livestock production, over the ten years after completion of the BPRE scheme.[57] This is 8.7 times the costs of the program, which suggests an extremely high benefit-cost ratio and implied (social) rate of return to the program.

We also estimate the cost-benefit for Wheat, Cotton and Livestock trainings separately. Table 22 provides the average cost per household for each training individually. Average cost per household for wheat trainings is PKR 23,351, while for cotton and livestock trainings they are PKR 24,820 and PKR 15,175 respectively. Considering the benefits from wheat trainings from the first year (2018) for average households in treatment villages, we obtain a total benefit

[56] The decay rate is calculated by taking the percent decrease in value of annual output from agri-livestock production from 2018 and 2019.

[57] These are the nominal values as we assumed zero inflation rates

of PKR 31,621, which is a 1.35 times return relative to the average cost per household of the wheat trainings. The additional benefits from wheat trainings in the second year come to about PKR 12,551.

Table 20 Cost-Benefit Breakdown for Wheat, Cotton, and Livestock

Training	Number of Trainees	Cost of Training (PKR)	Average Cost per Household (PKR)[+]	Projected 10 year Benefit (PKR)[++]
Wheat	10,509	213,385,286	23,351	52,427
Cotton	8,992	194,067,599	24,820	44,643
Livestock	12,175	160,657,271	15,175	49,567

Notes: [+]Average cost per household is calculated by dividing the cost of the training by the number of trainees for that training and multiplying it by 1.15 (average number of members per household trained is 1.15).
[++]Projected 10-year benefit is calculated using the respective constant decay rate of the benefits from each training.

Similarly, for cotton and livestock trainings, the benefits from each training from the first year (2018) for the average household in treatment villages are PKR 34,418 and PKR 27,601, respectively. This is about a 1.4 times return relative to the average cost per household of the cotton trainings, and a 1.8 times return relative to the average cost per household of the livestock trainings, considering only the benefits from the first year. The benefits from cotton and livestock trainings from the second year (2019) are PKR 7,884 and PKR 11,588, respectively, for the average household in treatment villages.

The projected 10 year benefits from wheat, cotton and livestock trainings come to about PKR 52,427, PKR 44,643, and PKR 49,567, respectively.[58]

[58] The decay rates used to calculate the projected benefits are 60.3%, 77.1%, 58% for wheat, cotton and livestock respectively.

10. Conclusion: Lessons Learnt and Recommendations

Our findings show that large-scale agri-livestock trainings can indeed have substantial impact and show extremely favourable benefit-cost ratios. This bodes well for scaling such programs up to a national level. The gains in production observed could have a first-order impact on an economy like Pakistan's given the importance of agriculture and livestock. The gains would not only alleviate food security concern, but the resulting reduction in imports could bode well for the current account balance, as Pakistan's cotton imports are on the rise.

However, our results also offer several words of caution and possible additional policy interventions. Interestingly, despite the improvement in knowledge and productivity immediately after training, we were unable to detect significant changes in practices. This suggests that either our measures of practices were not sufficiently accurate, or that perhaps the knowledge gained through trainings did not in fact change the measured practices, but rather increased the returns to their existing practices which resulted in increased output. This requires more exploration, especially if changed practices could further enhance productivity.

Second, our results show that linking agents across the agri-livestock value chain once through meetings and *melas* does not seem to have additional impact over the impact from training. This may either be because these linkages are already reasonable enough or that they matter less. From a policy perspective, this warrants further examination to see whether we need to design stronger linkage programs or that in fact these are not needed as the market naturally creates linkages as needed.

Finally, and perhaps most importantly, the impact of the training declined significantly over a one-year period. Intriguingly, this decline did not show up in milk yields, suggesting that different types of knowledge may show different levels of persistence, likely based on how regularly this knowledge is applied. Overall, this points to the concern that trainees tend to forget what they have learnt in the trainings and cannot sustain the knowledge gained from trainings and the resulting gains in production. It also points to the fact that knowledge retention for a seasonal farm activity (wheat and cotton) is harder as compared with a livestock management activity that continues all year round. This suggests that for one to continue to reap the benefits of the initial training, there needs to be a (low-cost) way of refresher trainings. To make such trainings cost-effective, we may need to send regular reminders of training course content regarding input usage and best farm practices to trainees through SMS or recorded voice calls (see Cole and Fernando, 2012; Larochelle et al., 2017; Casaburi et al., 2013).[59] Additionally, we may also try setting up call centres to provide constant feedback and tailored advice to farmers in accordance with their specific needs.

The broader message is that agri-livestock training—especially when provided at scale and through providers who actually "practice what they preach"—can be a powerful way of alleviating poverty, addressing food security and affecting the overall growth of an economy like Pakistan.

[59] Introduction of low-cost information and communications technology (ICT) has been shown to deliver timely, relevant, and actionable advice to farmers that can improve agricultural knowledge and yields in India and Kenya. See cited papers for more information.

References

Ahmed, S. and Gautam, A. (2013). Increasing Agricultural Productivity. The World Bank Group South Asia Region. Pakistan Policy Note 6.

Barnett, C., Masset, E. Dogbe, T., Jupp, D., Korboe, D., Acharya, A., Nelson, K., Eager, R. and Hilton, T. (2018). Impact Evaluation of the SADA Millennium Villages Project in Northern Ghana: Endline Summary Report. Brighton: Itad in association with IDS, LSHTM and PDA Ghana

Bedoya, G. (2019). Afghanistan Targeting the Ultra-Poor Impact Evaluation. AEA Randomized Controlled Trials. doi: 10.1257/rct.2665-2.0

Bonan, J. and Pagani, L. (2017). Junior Farmer Field Schools, Agricultural Knowledge and Spillover Effects: Quasi-Experimental Evidence from Northern Uganda. The Journal of Development Studies, 54(11), pp.2007-2022.

Bolwig, S., P. Gibbon, and S. Jones. 2009. The Economics of Smallholder Organic Contract Farming in Tropical Africa. World Development 37(6): 1094–1104.

Carpio, Ximena V. Del; Maredia, Mywish. (2011). Impact evaluations in agriculture: an assessment of the evidence (English). Washington, DC: World Bank. http://documents.worldbank.org/curated/en/131051468161681356/Impact-evaluations-in-agriculture-an-assessment-of-the-evidence

Casaburi, L., Kremer, M., Mullainathan, S. and Ramrattan, R. (2019 draft). *Harnessing ICT to Increase Agricultural Production: Evidence From Kenya* *. [online] Available at: https://scholar.harvard.edu/files/kremer/files/sms_paper_with_tables_20190923_merged.pdf

Cavatassi Romina, Lina Salazar, Mario Gonzalez-Flores, Paul Winters. (2011). How do Agricultural Programmes Alter Crop Production? Evidence from Ecuador. *Journal of Agricultural Economics* 62 (2): 403-428.

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Lodhi, A., Sheikh, S., Siddiqui, S., Tourek, G., Niazi, M., Shoaib, A. (2012). PEOP Household and Community Surveys: Baseline Household Report on Skills, *Center for Economic Research in Pakistan.*

Cheema, A., Khwaja, A., Naseer, F. & Shapiro, J. (2015). Big Push for the Rural Economy: Evaluation Design Proposal, Center for Economic Research in Pakistan.

Clemens, M. and Demombynes, G. (2013). The New Transparency in Development Economics: Lessons from the Millennium Villages Controversy. SSRN Electronic Journal.

Cole, S.A. and Fernando, A.N. (2012). The Value of Advice: Evidence from Mobile Phone-Based Agricultural Extension. SSRN Electronic Journal. [online] Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2179008

Cole, Shawn Allen and Fernando, A., 'Mobile'izing Agricultural Advice: Technology Adoption, Diffusion and Sustainability (April 29, 2016). Harvard Business School Finance Working Paper No. 13-047. Available at SSRN: https://ssrn.com/abstract=2179008 or http://dx.doi.org/10.2139/ssrn.2179008

Davis, K., Nkonya, E., Kato, E., Mekonnen, D., Odendo, M., Miiro, R. and Nkuba, J. (2012). Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa. World Development, 40(2), pp.402-413.

Feder, G., Murgai, R. and Quizon, J. (2003). Sending Farmers Back to School: The Impact of Farmer Field Schools in Indonesia. Policy Research Working Papers.

Guo, M., Jia, X., Huang, J., Kumar, K., & Burger, N. (2015). Farmer field school and farmer knowledge acquisition in rice production: Experimental evaluation in China. Agriculture, Ecosystems & Environment, 209, 100-107. doi: 10.1016/j.agee.2015.02.011

Habib, Mubashir & Zafarullah Khan, Muhammad & Iqbal, Mahmood & Nawab, Khalid & Ali, Shahid. (2007). Role of farmer field schools on sugarcane productivity in Malakand, Pakistan. African Crop Science Conference Proceedings.

IEG (Independent Evaluation Group). (2012). World Bank Group Impact Evaluations: Relevance and Effectiveness. Washington, DC: World Bank Group.

Khan, M. and Iqbal, M. (2005). Sustainable Cotton Production through Skill Development among Farmers: Evidence from Khairpur District of Sindh, Pakistan. *The Pakistan Development Review*, 44(4II), pp.695-716.

Khatam, A., Sher, M., & Ashraf, I. (2013). Role of farmer field schools in enhancing skills of farming community in Khyber Pakhtunkhwa, Pakistan. *Pakistan Journal of Agricultural Research*, 26(1), 59-64.

Kremer, M. (1993). Population growth and technological change: one million BC to 1990. *The Quarterly Journal of Economics*, 681-716.

Larochelle, C., Alwang, J., Travis, E., Barrera, V.H. and Dominguez Andrade, J.M. (2017). Did You Really Get the Message? Using Text Reminders to Stimulate Adoption of Agricultural Technologies. *The Journal of Development Studies*, 55(4), pp.548–564.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of development Economics*, 99(2), 210–221.

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1989). Industrialization and the Big Push. *Journal of Political Economy, 97*(5), 1003-1026.

Nankhuni, F., & Paniagua, G. (2013). Meta-evaluation of Private Sector Interventions in Agribusiness. *Washington DC: World Bank*.

Nziguheba, G., Palm, C., Berhe, T., Denning, G., Dicko, A., Diouf, O., Diru, W., Flor, R., Frimpong, F., Harawa, R., Kaya, B., Manumbu, E., McArthur, J., Mutuo, P., Ndiaye, M., Niang, A., Nkhoma, P., Nyadzi, G., Sachs, J., Sullivan, C., Teklu, G., Tobe, L. and Sanchez, P. (2010). The African Green Revolution. *Advances in Agronomy*, pp.75-115.

Quizon, J., Feder, G. and Murgai, R. (2001). Fiscal Sustainability of Agricultural Extension: The Case of the Farmer Field School Approach. *Journal of International Agricultural and Extension Education*, 8(1).

Rasul, I., Khan, A., Gondal, O., Sheikh, S., Shoaib, A., Siddiqui, S. (2012). PEOP Household and Community Surveys: Baseline Household Report on Livestock, *Centre for Economic Research in Pakistan*.

Rejesus, R., Mutuc, M., Yasar, M., Lapitan, A., Palis, F. and Chi, T. (2012). Sending Vietnamese Rice Farmers Back to School: Further Evidence on the Impacts of Farmer

Field Schools. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 60(3), pp.407-426.

Rosenstein-Rodan, P. (1943). Problems of Industrialisation of Eastern and South-Eastern Europe. *The Economic Journal,53*(210/211), 202-211. doi:10.2307/222631

Sachs, J. (1999). *Millennium Villages*. *Millenniumvillages.org*. Retrieved 18 April 2016, from http://millenniumvillages.org/

Siddiqui, A.A., Z.A. Mirani and S.S. Bukhari. (2012). Impact of Farmer Field School training on farmers' knowledge and skills regarding recommended agro-ecological sound IPM practices in selected districts of Sindh. *Pak. J. Agric., Agric. Eng.* Vet. Sci. 28: 186-198.

Todo, Y. and Takahashi, R. (2011). Impact of Farmer Field Schools on Agricultural Income and Skills: Evidence from an Aid-Funded Project in Rural Ethiopia. *Journal of International Development*, 25(3), pp.362-381.

Van Breukelen, G. J. (2006). Ancova versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of clinical epidemiology*, 59(9), 920–925.

Waddington, H., Snilstveit, B., Hombrados, J., Vojtkova, M., Phillips, D., Davies, P. and White, H. (2014). Farmer Field Schools for Improving Farming Practices and Farmer Outcomes: A Systematic Review. Campbell Systematic Reviews, 10(1), p.i-335.

Wanjala, B. and Muradian, R. (2013). Can Big Push Interventions Take Small-Scale Farmers out of Poverty? Insights from the Sauri Millennium Village in Kenya. World Development, 45, pp.147-160.

World Bank. 2012. World Bank Group Agriculture Action Plan 2013 -2015. Draft July 12, 2012. Washington, DC: World Bank.

Appendix A: Details on Training Courses

Table A.1 Details on Wheat Course

BASIC FARMER TRAINING	**Wheat and General Agriculture Training**			
	Wheat Module			
		Module Name	**Description of the module/What will be taught in module**	**Duration**
	1.	Seed Selection and Quality	• How to assess seed quality by outward characteristics? • What seed varieties give better yields and should be used? • How often should the famers change seed varieties? • Variety specific fertilizer requirement? Variety specific disease susceptibility?	2 days , 50% theory and 50% practical
	2.	Land Preparation (Tillage, Land leveling)	• How to level land by traditional methods and/or by laser-leveling? Training on land preparation (tillage etc.)	2 days , 30% theory and 70% practical
	3.	Planting (timing, seed priming)	• What is appropriate time to sow seeds? • Training on planting method for better results • Seed Priming Seed treatment	3 days , 30% theory and 70% practical
	4.	Fertilizer and Pesticide (quantity, type, frequency of application)	• What varieties of fertilizer/pesticide should be used? • What quantity of fertilizer/pesticide should be applied? At what intervals should the fertilizer/pesticide should be applied?	10 days , 30% theory and 70% practical
	5.	Water Management (timing, quantity, frequency)	• What quantity of water is to be given to the crop at different stages of crop growth? Training on water conservation methods.	10 days , 30% theory and 70% practical
	6.	Post-Harvest (thrashing, storage, transportation)	• How to reduce crop wastage at the time of harvest? After crop is harvested, how to store/transport the crop to reduce/minimize crop loss?	5 days , 30% theory and 70% practical
	General Agriculture			
	1.	Soil Management (Soil testing, manure, compost, fallow)	• How to assess the soil quality by looking visible characteristics? • How can one have the soil tested in a laboratory? • Depending on the soil quality, when and how much manure/fertilizer/compost should normally be added. • How often should the land be kept fallow and at what intervals? How to make and apply composts and manures?	3 days , 50% theory and 50% practical
	2.	Record keeping	• Keeping paper-based records of costs of inputs (seeds, fertilizer etc.), labor (farm wages) and overheads (electricity, fuel, and other costs) • Keeping paper-based records of revenues (from sales of the produce) Calculating profit or loss for each season	1 days , 50% theory and 50% practical

	3.	Market (sale price negotiation)	• How to know the prevailing market/mandi rates for the produce? How to bargain with customers/middlemen for better rates?	2 days , 60% theory and 40% practical
	4.	Farm Management	• How many farm labor one should employ given a piece of land? Sources of formal loans/credit that farmers can use?	1 days , 100% theory
	5.	Machinery Maintenance and urgent repairs	• Farm machinery used at small to medium farm in cotton-wheat system • Routine maintenance of farm machinery • Urgent on farm repairs of machinery When to seek 'Workshop' help for repair	3 days: 30% theory and 70% practical
	6.	Fodder Module	• Seed Selection and Quality • Land Preparation (Tillage, Land leveling) • Planting (timing, seed priming) • Fertilizer and Pesticide (quantity, type, frequency of application) • Water Management (timing, quantity, frequency) • Post-Harvest (thrashing, storage, transportation) • Personal Protection & Safety	10 days

Table A.2 Details on Cotton Course

BASIC FARMER TRAINING				
Cotton and General Agriculture Training				
Cotton Module				
		Module Name	**Description of the module/What will be taught in module**	**Duration**
	1.	Seed Selection and Quality	• How to assess seed quality by outward characteristics? • What seed varieties give better yields and should be used? • How often should the famers change seed varieties? • Variety specific fertilizer requirement? Variety specific disease susceptibility?	3 days , 70% theory and 30% practical
	2.	Land Preparation (Tillage, Land leveling)	• How to level land by traditional methods and/or by laser-leveling? Training on land preparation (tillage etc.)	3 days , 30% theory and 70% practical
	3.	Planting (timing, seed priming)	• What is appropriate time to sow seeds? • Training on planting methods for better results? Seed treatment	3 days , 30% theory and 70% practical
	4.	Fertilizer and Pesticide (quantity, type, frequency of application)	• What varieties of fertilizer/pesticide should be used? • What quantity of fertilizer/pesticide should be applied? At what intervals should the fertilizer/pesticide should be applied?	2 days , 30% theory and 70% practical
	5.	Water Management (timing, quantity, frequency)	• What quantity of water is to be given to the crop at different stages of crop growth? • Training on water conservation methods. Water channel management	15 days , 30% theory and 70% practical
	6.	Post-Harvest (thrashing, storage, transportation)	• How to reduce crop wastage at the time of harvest? After crop is harvested, how to store/transport the crop to reduce/minimize crop loss?	5 days , 30% theory and 70% practical
	General Agriculture			
	1.	Soil Management (Soil testing, manure, compost, fallow)	• How to assess the soil quality by looking visible characteristics? • How can one have the soil tested in a laboratory? • Depending on the soil quality, when and how much manure/fertilizer/compost should normally be added. • How often should the land be kept fallow and at what intervals? How to make and apply composts and manures?	3 days , 50% theory and 50% practical
	2.	Record keeping	• Keeping paper-based records of costs of inputs (seeds, fertilizer etc.), labor (farm wages) and overheads (electricity, fuel, and other costs) • Keeping paper-based records of revenues (from sales of the produce) Calculating profit or loss for each season	1 days , 50% theory and 50% practical
	3.	Market (sale price negotiation)	• How to know the prevailing market/mandi rates for the produce?	2 days , 60% theory and 40% practical

			How to bargain with customers/middlemen for better rates?	
	4.	Farm Management	• How many farm labor one should employ given a piece of land? Sources of formal loans/credit that farmers can use?	1 days , 100% theory
	5.	Machinery Maintenance and urgent repairs	• Farm machinery used at small to medium farm in cotton-wheat system • Routine maintenance of farm machinery • Urgent on farm repairs of machinery When to seek 'Workshop' help for repair	3 days: 30% theory and 70% practical
	6.	Fodder Module	• Seed Selection and Quality • Land Preparation (Tillage, Land leveling) • Planting (timing, seed priming) • Fertilizer and Pesticide (quantity, type, frequency of application) • Water Management (timing, quantity, frequency) • Post-Harvest (thrashing, storage, transportation) • Personal Protection & Safety	10 days

Table A.3 Details on Livestock Course

Basic Livestock Training (14 days)		
Modules	**Details**	**Duration**
Animal Health	Immunity and vaccination	2 days , 60 percent time on theory and 40 percent practical
	Deworming	
	Mastitis Control	
	Ticks Management	
	Hygiene and Bio Security	
	Protocols to follow for vaccination etc.	
	Major diseases	
	Outbreak management	
	Stress Management	
Feed and Nutrition	Animal requirement of feed	3 days , 70 percent time on theory and 30 percent practical
	Introduction to DM based feeding and its calculations	
	Daily protein and energy requirements	
	Different fodders and their nutritional value	
	Requirements of Concentrates & Minerals	
	TMR based feeding	
	Feed evaluations	
	Feed storage	
	Importance of free access to water	
Fodder Preservation	Fodder preservation techniques & importance.	2 day 40 percent theory and 60 % practical
	Different machinery available.	
Milking Protocols and Handling	Milk let down procedures.	2 day 60 percent theory and 40 % practical
	Hormones involved.	
	Milk let down stimuli.	
	Importance of regular interval.	
	Consequences of irregular intervals.	
	Benefits of regular intervals and how it is achieved.	
	Hand Milking procedures: pre-dipping, post-dipping etc.	
	CIP protocols, Colostrum management etc.	
Breed Selection and Improvement	Different Breeds of Cows & Buffalo.	3 days , 50 percent time on theory and 50 percent practical
	Exotic breeds.	
	Characteristics of breeds.	
	Linear scoring.	
	Introduction to AI.	
	Getting cows pregnant ASAP after parturition.	
	Nine/zero model.	
	Anatomy of reproductive organs	
Farm Management	Livestock farm economics.	2 day 70 percent theory and 30 percent practical
	Methods of record keeping of animal herd.	
	Farm input Vs. Farm Out Put.	
	How to maximize profit.	
	Cost/benefit analysis	

Table A.4 BPRE trainings rollout timeline

Training	**Timeline**
Wheat	December 2016 – April 2017
Cotton	May 2017 – November 2017
Kitchen Gardening	September 2017 – January 2018
Basic Livestock	October 2017 – March 2018
Specialised Agricultural	
Farm Machinery Mechanic	September 2017 – March 2018
Electrician	September 2017 – April 2018
Specialised Livestock	
Village Milk Collector	April 2017 – July 2018
Animal Health Worker	April 2017 – July 2018
Artificial Insemination Technician	August 2017 – August 2018
Farm Supervisor	July 2017 – March 2018
Village melas (linkage component)	**April – August 2018**

Appendix B: Enrolment in Training Courses

The census activity was conducted at the end of the training courses and served as a check on the attendance data. The purpose of the census activity was to collect basic data on trainees, identifying trainees from our BPRE sample households, regardless of whether they had been provided with vouchers. This exercise was conducted to also enable for the inclusion of trainees from our BPRE ineligible sample (i.e. trainees that were members of BPRE sample households but not recorded as participating in agricultural activities and hence not provided with vouchers). The types of enrolees in the training courses identified during the census activity were:

- **BPRE eligible**: BPRE sample households that were distributed a voucher for the course
- **BPRE ineligible**: BPRE sample households that were not distributed a voucher for the course
- **Non-BPRE** (in-village): Individuals / households that are not in the BPRE sample but residents of the village where the training is being conducted (and are therefore likely to benefit the village economy)
- **Out-of-village**: Individuals / households that are not resident in the village where the training is being conducted

Wheat course	**BPRE eligible**	**BPRE ineligible**	**Non – BPRE (in-village)**	**Total in-village**	**Out-of-village**	**Total**
Individual (n)	2,209	492	3,695	6,396	1,489	7,885
Household (n)	1,981	448	3,223	5,652	1,320	6,972
Household (%)	43%	12%	20%	23%	N/A	N/A

Cotton course	BPRE eligible	BPRE ineligible	Non – BPRE (in-village)	Total in-village	Out-of-village	Total
Individual (n)	1,902	521	3,833	6,256	1,792	8,048
Household (n)	1,686	462	3,416	5,564	1,666	7,230
Household (%)	36%	12%	21%	22%	N/A	N/A

Basic Livestock Course	BPRE eligible	BPRE ineligible	Non – BPRE (in-village)	Total in-village	Out-of-village	Total
Phase 1						
Number of Households	1,178	312	2,215	3,703	490	4,193
Phase 2						
Number of Households	968	255	1,801	3,024	144	3,168
Phase 3						
Number of Households	571	146	1,227	1,944	140	2,084

Appendix C: Balance Checks

		T (1 & 2)	T1	T2	C	P-value T vs C	P-value T1 vs T2
HH Controls	Household head is literate	0.976	0.982	0.970	0.974	0.788	0.199
		(0.006)	(0.006)	(0.009)	(0.006)		
	Family size	6.134	6.073	6.196	6.047	0.181	0.305
		(0.064)	(0.096)	(0.085)	(0.091)		
	Age of household head	46.964	46.870	47.059	47.385	0.493	0.855
		(0.432)	(0.597)	(0.634)	(0.529)		
	Agriculture asset index (PCA)	-0.055	-0.066	-0.044	0.114	0.003	0.708
		(0.034)	(0.044)	(0.054)	(0.062)		
	Non-business asset index (PCA)	0.015	-0.055	0.086	-0.032	0.621	0.388
		(0.110)	(0.145)	(0.169)	(0.166)		
	Own or have access to a tubewell (0/1)	1.912	1.868	1.960	1.863	0.289	0.084
		(0.037)	(0.038)	(0.063)	(0.045)		
Land Features	Land suffers waterlogging (y=1 n=2)	1.782	1.831	1.728	1.781	0.963	0.065
		(0.034)	(0.042)	(0.050)	(0.043)		
	Land suffers salinity (y=1 n=2)	1.806	1.844	1.765	1.825	0.596	0.104
		(0.029)	(0.037)	(0.043)	(0.031)		
Agricultural Production	Grew crops one year prior to 2016 survey (0/1)	0.345	0.352	0.338	0.425	0.001	0.780
		(0.021)	(0.031)	(0.029)	(0.018)		
	Grew wheat one year prior to 2016 survey (0/1)	0.280	0.284	0.275	0.342	0.010	0.902
		(0.021)	(0.029)	(0.031)	(0.019)		
	Grew cotton one year prior to 2016 survey (0/1)	0.192	0.202	0.181	0.226	0.201	0.277
		(0.018)	(0.026)	(0.026)	(0.029)		
	Last year's wheat output (maund)	37.208	35.360	39.081	51.574	0.019	0.444
		(3.788)	(4.662)	(6.061)	(6.646)		
	Last year's cotton output (maund)	12.542	12.690	12.391	17.039	0.188	0.797
		(1.597)	(2.121)	(2.427)	(3.852)		
Agri Knowledge and Practice	Years of experience with growing wheat	17.474	16.993	17.978	16.958	0.437	0.485
		(0.723)	(1.026)	(1.002)	(0.755)		
	Years of experience with growing cotton	18.039	17.527	18.618	17.640	0.480	0.650
		(0.705)	(1.025)	(0.929)	(0.824)		
	Additive agricultural knowledge index	0.519	0.518	0.520	0.516	0.848	0.922
		(0.007)	(0.008)	(0.012)	(0.010)		

	Additive agricultural practice index	0.255	0.249	0.262	0.260	0.599	0.592
		(0.011)	(0.013)	(0.018)	(0.017)		
Livestock Production	Produced milk one year prior to 2016 survey (0/1)	0.421	0.423	0.419	0.485	0.021	0.971
		(0.017)	(0.023)	(0.025)	(0.023)		
	Daily milk output per cow (liter)	3.943	3.895	3.992	4.130	0.207	0.511
		(0.080)	(0.124)	(0.104)	(0.155)		
	Daily milk output per buffalo (liter)	5.173	5.126	5.223	5.150	0.576	0.817
		(0.134)	(0.218)	(0.152)	(0.206)		
Livestock Knowledge	Total number of adult animals owned	0.922	0.943	0.900	1.122	0.024	0.625
		(0.058)	(0.074)	(0.089)	(0.089)		
	Grew fodder one year prior to 2016 survey (0/1)	0.261	0.274	0.247	0.313	0.027	0.369
		(0.018)	(0.026)	(0.024)	(0.017)		
	Livestock knowledge index	0.498	0.505	0.492	0.477	0.017	0.113
		(0.007)	(0.010)	(0.011)	(0.008)		
	Livestock practice index	0.246	0.248	0.245	0.255	0.328	0.660
		(0.006)	(0.008)	(0.009)	(0.009)		

Appendix D: List of Covariates

Table C.1 List of covariates used

Covariates used in Regressions	
Wheat Controls	Household Size
	Asset Index
	Agriculture Index
	Household Head Literacy
	Fertility of Land
	Land suffers from salinity (Dummy)
	Land suffers from waterlogging (Dummy)
	Soil Quality
	Access to tube well
	Quantity of fertilizer used in wheat production
	Number of years for which household has grown wheat
	Land cultivated for agriculture
	Non cultivated land owned by household
Cotton Controls	Household Size
	Asset Index
	Agriculture Index
	Household Head Literacy
	Land Fertility
	Land suffers from salinity (Dummy)
	Land suffers from waterlogging (Dummy)
	Soil Quality
	Access to tube well
	Quantity of fertilizer used in cotton production
	Number of years for which household has grown cotton
	Land cultivated for agriculture
	Non cultivated land owned by household
Livestock Controls	Household Size
	Asset Index
	Household Head Literacy
	Household grew fodder
	Number of Livestock owned by household

Appendix E: Individual Income regressions

1. *Wheat*

	Ln of Wheat Production Worth		
	(1) ITT	(2) ITT	(3) LATE
Panel A: Tracker 2018			
Treated	0.715***	0.797***	
	(0.226)	(0.234)	
Enrolled			1.459***
			(0.448)
Ineligible for Any Training		-1.188***	
		(0.208)	
Treated × Ineligible		-0.164	
		(0.266)	
Baseline 2016	0.508***	0.490***	0.495***
	(0.015)	(0.015)	(0.015)
Obs.	11351	11351	11351
Covariates	X	X	X
Impact on Ineligible		0.6334	
(p-value)		0.0151	
Panel B: Endline 2019			
Treated	0.350*	0.399*	
	(0.200)	(0.210)	
Enrolled			0.714*
			(0.401)
Ineligible for Any Training		-1.141***	
		(0.215)	
Treated × Ineligible		-0.124	
		(0.260)	
Baseline 2016	0.477***	0.460***	0.470***
	(0.015)	(0.015)	(0.015)
Obs.	11133	11133	11133
Covariates	X	X	X
Impact on Ineligible		0.2749	
(p-value)		0.2319	
Panel C: Cross-Year Diff 18-19			
Diff in ITT Effect	-0.364**		
	(0.147)		
Diff in ITT Effect for Eligible		-0.398**	
		(0.168)	
Diff in LATE			-0.745**
			(0.294)
Obs.	22484	22484	22484
Covariates	X	X	X
Diff in ITT Effect for Ineligible		-0.3584	
(p-value)		0.0763	

Notes: Regressions of log of wheat production worth on treatment. Standard errors clustered at the village level reported in parentheses.

2. Cotton

	Ln of Cotton Production Worth		
	(1) ITT	(2) ITT	(3) LATE
Panel A: Tracker 2018			
Treated	1.008***	1.160***	
	(0.248)	(0.282)	
Enrolled			2.067***
			(0.492)
Ineligible for Any Training		-0.665**	
		(0.261)	
Treated × Ineligible		-0.593*	
		(0.343)	
Baseline 2016	0.479***	0.470***	0.461***
	(0.018)	(0.019)	(0.019)
Obs.	11351	11351	11351
Covariates	X	X	X
Impact on Ineligible		0.5674	
(p-value)		0.0425	
Panel B: Endline 2019			
Treated	0.607***	0.705***	
	(0.187)	(0.223)	
Enrolled			1.243***
			(0.376)
Ineligible for Any Training		-0.680***	
		(0.223)	
Treated × Ineligible		-0.443	
		(0.312)	
Baseline 2016	0.452***	0.444***	0.441***
	(0.016)	(0.016)	(0.017)
Obs.	11133	11133	11133
Covariates	X	X	X
Impact on Ineligible		0.2616	
(p-value)		0.2534	
Panel C: Cross-Year Diff 18-19			
Diff in ITT Effect	-0.401**		
	(0.175)		
Diff in ITT Effect for Eligible		-0.456**	
		(0.201)	
Diff in LATE			-0.824**
			(0.349)
Obs.	22484	22484	22484
Covariates	X	X	X
Diff in ITT Effect for Ineligible		-0.3058	
(p-value)		0.0878	

Notes: Regressions of log of cotton production worth on treatment. Standard errors clustered at the village level reported in parentheses.

	Ln of Milk Production Worth		
	(1) ITT	(2) ITT	(3) LATE
Panel A: Tracker 2018			
Treated	0.490**	0.542***	
	(0.204)	(0.205)	
Enrolled			1.005**
			(0.413)
Ineligible for Any Training		-1.319***	
		(0.223)	
Treated × Ineligible		-0.018	
		(0.268)	
Baseline 2016	0.145***	0.129***	0.143***
	(0.020)	(0.020)	(0.020)
Obs.	11351	11351	11351
Covariates	X	X	X
Impact on Ineligible		0.5233	
(p-value)		0.0503	
Panel B: Endline 2019			
Treated	0.033	0.178	
	(0.180)	(0.153)	
Enrolled			0.067
			(0.366)
Ineligible for Any Training		-1.740***	
		(0.276)	
Treated × Ineligible		-0.142	
		(0.331)	
Baseline 2016	0.190***	0.168***	0.190***
	(0.015)	(0.014)	(0.015)
Obs.	11128	11128	11128
Covariates	X	X	X
Impact on Ineligible		0.0363	
(p-value)		0.9193	
Panel C: Cross-Year Diff 18-19			
Diff in ITT Effect	-0.458***		
	(0.160)		
Diff in ITT Effect for Eligible		-0.364*	
		(0.185)	
Diff in LATE			-0.939***
			(0.324)
Obs.	22479	22479	22479
Covariates	X	X	X
Diff in ITT Effect for Ineligible		-0.4870	
(p-value)		0.0173	

Notes: Regressions of log of milk production worth on treatment. Standard errors clustered at the village level reported in parentheses.

Appendix F: Analysis on Specialised Service Providers

1. Availability

	Availability					
	Agriculture Related Service Providers		Livestock Related Service Providers		Combined	
	(1) ITT	(2) ITT	(1) ITT	(2) ITT	(1) ITT	(2) ITT
Treated	0.0506		0.0301		0.0331	
	(0.0534)		(0.0451)		(0.0422)	
Eligible		0.0251		0.0129		0.0215
		(0.0549)		(0.0436)		(0.0431)
Ineligible		0.0789		0.0597		0.0632
		(0.0565)		(0.0581)		(0.0514)
Obs.	10,997	10,997	10,997	10,997	10,997	10,997

2. Accessibility

	Accessibility					
	Agriculture Related Service Providers		Livestock Related Service Providers		Combined	
	(1) ITT	(2) ITT	(1) ITT	(2) ITT	(1) ITT	(2) ITT
Treated	0.0486		0.0269		0.0288	
	(0.0498)		(0.0456)		(0.04)	
Eligible		0.0167		-0.0002		0.0142
		(0.0527)		(0.0456)		(0.0409)
Ineligible		0.089*		0.0748		0.0723
		(0.0535)		(0.0561)		(0.0476)
Obs.	10,997	10,997	10,997	10,997	10,997	10,997

3. Quality

	Quality					
	Agriculture Related Service Providers		Livestock Related Service Providers		Combined	
	(1) ITT	(2) ITT	(1) ITT	(2) ITT	(1) ITT	(2) ITT
Treated	-0.1167		-0.063		-0.0798	
	(0.0739)		(0.0622)		(0.0615)	
Eligible		-0.1319		-0.0814		-0.0885
		(0.0772)		(0.0595)		(0.0616)
Ineligible		-0.0707		-0.0313		-0.0724
		(0.0776)		(0.0758)		(0.0692)
Obs.	9,011	9,011	10,229	10,229	10,688	10,688

Appendix G: Comparison between T1 and T2 Villages

1. *Wheat 2019*

	Ln of Quantity Produced			Extensive Margin			Yield (Maund per Kanal)		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE	(7) ITT	(8) ITT	(9) LATE
Endline 2019									
T1	0.206**	0.250***		0.049**	0.060***		0.030	0.050	
	(0.082)	(0.092)		(0.019)	(0.021)		(0.048)	(0.071)	
T2	0.107	0.211**		0.020	0.044**		0.048	0.036	
	(0.083)	(0.092)		(0.019)	(0.020)		(0.048)	(0.071)	
Enrolled in T1			0.609***			0.144***			0.086
			(0.235)			(0.054)			(0.141)
Enrolled in T2			0.357			0.066			0.164
			(0.278)			(0.065)			(0.163)
Ineligible for Agri Training		-0.290***			-0.072***			0.031	
		(0.087)			(0.021)			(0.055)	
T1 × Ineligible		-0.149			-0.032			-0.044	
		(0.113)			(0.026)			(0.068)	
T2 × Ineligible		-0.231**			-0.050**			-0.004	
		(0.106)			(0.025)			(0.065)	
Baseline 2013	0.423***	0.421***	0.424***	0.273***	0.261***	0.258***	0.103*	0.103*	0.103*
	(0.047)	(0.046)	(0.047)	(0.052)	(0.052)	(0.052)	(0.058)	(0.058)	(0.058)
Baseline 2016	0.177***	0.173***	0.170***	0.177***	0.173***	0.171***	0.000	0.000	0.000
	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.000)	(0.000)	(0.000)
Intercept	0.468	0.491	0.438	0.177	0.194	0.192	3.130***	3.137***	3.082***
	(0.529)	(0.523)	(0.529)	(0.125)	(0.124)	(0.124)	(0.445)	(0.446)	(0.450)
Obs.	11133	11133	11133	11133	11133	11133	11133	11133	11133
Covariates	X	X	X	X	X	X	X	X	X
T2 - T1	-0.0986	-0.0389	-0.2518	-0.0288	-0.0158	-0.0785	0.0185	-0.0137	0.0773
(p-value)	0.1943	0.6736	0.2948	0.1060	0.4418	0.1643	0.6533	0.8233	0.5684
T1 Impact on Ineligible		0.1008			0.0281			0.0060	
(p-value)		0.3349			0.2589			0.8820	
T2 Impact on Ineligible		-0.0202			-0.0062			0.0321	
(p-value)		0.8371			0.7950			0.3766	

Notes: Regressions of crop production variables on T1 and T2. Standard errors clustered at the village level reported in parentheses.

2. Cotton 2019

	Ln of Quantity Produced			Extensive Margin			Yield (Maund per Kanal)		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE	(7) ITT	(8) ITT	(9) LATE
Endline 2019									
T1	0.181***	0.201**		0.064***	0.077***		-0.030	-0.057	
	(0.063)	(0.092)		(0.016)	(0.023)		(0.038)	(0.062)	
T2	0.144**	0.263***		0.048***	0.082***		-0.009	-0.014	
	(0.072)	(0.097)		(0.017)	(0.023)		(0.041)	(0.066)	
Enrolled in T1			0.647***			0.229***			-0.108
			(0.227)			(0.057)			(0.134)
Enrolled in T2			0.556**			0.186***			-0.032
			(0.273)			(0.066)			(0.160)
Ineligible for Agri Training		-0.169**			-0.045**			-0.038	
		(0.074)			(0.019)			(0.049)	
T1 × Ineligible		-0.089			-0.041			0.066	
		(0.115)			(0.030)			(0.074)	
T2 × Ineligible		-0.263**			-0.074**			0.013	
		(0.113)			(0.029)			(0.075)	
Baseline 2013	0.377***	0.375***	0.374***	0.146***	0.144***	0.139***	0.037*	0.037*	0.037*
	(0.041)	(0.042)	(0.041)	(0.037)	(0.038)	(0.038)	(0.021)	(0.021)	(0.020)
Baseline 2016	0.165***	0.161***	0.156***	0.168***	0.163***	0.154***	0.001	0.001	0.001
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.001)	(0.001)	(0.001)
Intercept	0.491	0.509	0.422	0.444***	0.449***	0.426***	1.387***	1.394***	1.394***
	(0.496)	(0.493)	(0.488)	(0.131)	(0.131)	(0.126)	(0.292)	(0.298)	(0.295)
Obs.	11133	11133	11133	11133	11133	11133	11133	11133	11133
Covariates	X	X	X	X	X	X	X	X	X
T2 - T1	-0.0366	0.0629	-0.0910	-0.0158	0.0049	-0.0435	0.0215	0.0430	0.0764
(p-value)	0.6179	0.5201	0.7388	0.3707	0.8378	0.5112	0.5841	0.4896	0.6082
T1 Impact on Ineligible		0.1117			0.0362			0.0089	
(p-value)		0.1219			0.0650			0.7986	
T2 Impact on Ineligible		0.0002			0.0082			-0.0013	
(p-value)		0.9984			0.6924			0.9735	

Notes: Regressions of crop production variables on T1 and T2. Standard errors clustered at the village level reported in parentheses.

3. *Milk 2019*

	Ln of Quantity Produced			Extensive Margin			Yield (Output per Animal)		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE	(7) ITT	(8) ITT	(9) LATE
Endline 2019									
T1	0.060*	0.100***		0.007	0.013		0.208***	0.268***	
	(0.033)	(0.034)		(0.020)	(0.019)		(0.057)	(0.074)	
T2	0.063*	0.096***		0.008	0.016		0.213***	0.289***	
	(0.033)	(0.032)		(0.020)	(0.017)		(0.055)	(0.071)	
Enrolled in T1			0.177*			0.019			0.613***
			(0.096)			(0.058)			(0.173)
Enrolled in T2			0.192*			0.025			0.648***
			(0.099)			(0.061)			(0.167)
Ineligible for LVS Training		-0.334***			-0.254***			0.228***	
		(0.044)			(0.033)			(0.064)	
T1 × Ineligible		-0.047			0.023			-0.212**	
		(0.049)			(0.034)			(0.083)	
T2 × Ineligible		-0.059			0.003			-0.244***	
		(0.060)			(0.038)			(0.086)	
Baseline 2013	0.310***	0.227***	0.303***	0.292***	0.185***	0.291***	0.129***	0.129***	0.127***
	(0.018)	(0.017)	(0.018)	(0.013)	(0.014)	(0.013)	(0.015)	(0.015)	(0.015)
Baseline 2016	0.000	0.000	0.000	0.205***	0.177***	0.205***	0.061***	0.062***	0.062***
	(.)	(.)	(.)	(0.011)	(0.011)	(0.011)	(0.008)	(0.008)	(0.008)
Intercept	0.253***	0.444***	0.257***	0.206***	0.357***	0.207***	2.730***	2.673***	2.698***
	(0.082)	(0.082)	(0.083)	(0.040)	(0.043)	(0.040)	(0.114)	(0.118)	(0.111)
Obs.	11132	11132	11132	11132	11132	11132	11132	11132	11132
Covariates	X	X	X	X	X	X	X	X	X
T2 - T1	0.0027	-0.0037	0.0146	0.0016	0.0034	0.0055	0.0044	0.0210	0.0347
(p-value)	0.9313	0.9168	0.8758	0.9243	0.8276	0.9116	0.9331	0.7675	0.8286
T1 Impact on Ineligible		0.0531			0.0354			0.0558	
(p-value)		0.2629			0.2810			0.3118	
T2 Impact on Ineligible		0.0372			0.0186			0.0446	
(p-value)		0.5127			0.6143			0.4593	

Notes: Regressions of milk production variables on T1 and T2. Standard errors clustered at the village level reported in parentheses.

4. Knowledge 2019 – Producers only

	Agri Knowledge		Lvs Knowledge	
	(1) ITT	(2) LATE	(3) ITT	(4) LATE
Endline 2019				
T1	-0.000		0.008*	
	(0.007)		(0.005)	
T2	-0.002		0.009**	
	(0.008)		(0.005)	
Enrolled in T1		-0.000		0.021*
		(0.015)		(0.012)
Enrolled in T2		-0.005		0.026**
		(0.018)		(0.012)
Baseline 2016	0.011	0.011	0.008	0.008
	(0.013)	(0.013)	(0.007)	(0.007)
Intercept	0.479***	0.480***	0.560***	0.560***
	(0.051)	(0.052)	(0.011)	(0.010)
Obs.	7070	7070	8685	8685
Covariates	X	X	X	X
T2 - T1	.	.	.	.
(p-value)	0.7458	0.7332	0.7464	0.6556

Notes: Regressions of knowledge indices (producers only) on T1 and T2. Standard errors clustered at the village level reported in parentheses.

5. *Knowledge 2019 – General*

	Agri Knowledge			Lvs Knowledge		
	(1) ITT	(2) ITT	(3) LATE	(4) ITT	(5) ITT	(6) LATE
Endline 2019						
T1	0.005	0.008		0.009	0.006	
	(0.010)	(0.009)		(0.005)	(0.005)	
T2	0.016	0.005		0.008	0.009*	
	(0.011)	(0.011)		(0.005)	(0.005)	
Enrolled in T1			0.013			0.024
			(0.026)			(0.015)
Enrolled in T2			0.048			0.023
			(0.033)			(0.015)
Ineligible		-0.025**			-0.014*	
		(0.011)			(0.008)	
1.treated1.ineligible		-0.015			0.006	
		(0.014)			(0.007)	
2.treated1.ineligible		0.023*			-0.004	
		(0.013)			(0.007)	
Baseline 2016	0.016	0.013	0.016	0.003	0.003	0.003
	(0.017)	(0.017)	(0.017)	(0.010)	(0.010)	(0.010)
Intercept	0.654***	0.666***	0.641***	0.574***	0.575***	0.573***
	(0.067)	(0.067)	(0.070)	(0.015)	(0.015)	(0.015)
Obs.	11133	11133	11133	11132	11132	11132
Covariates	X	X	X	X	X	X
T2 - T1	0.0109	-0.0028	0.0355	-0.0010	0.0031	-0.0016
(p-value)	0.3240	0.7993	0.2671	0.8318	0.5440	0.9060
T1 Impact on Ineligible		-0.0068			0.0115	
(p-value)		0.6498			0.1498	
T2 Impact on Ineligible		0.0276			0.0051	
(p-value)		0.0667			0.5398	

Notes: Regressions of knowledge indices (general) on T1 and T2. Standard errors clustered at the village level reported in parentheses.

6. *Practice 2019*

	Agri Practice		Lvs Practice	
	(1) ITT	(2) LATE	(3) ITT	(4) LATE
Endline 2019				
T1	-0.001		0.004	
	(0.012)		(0.012)	
T2	-0.004		-0.016	
	(0.012)		(0.011)	
Enrolled in T1		-0.002		0.011
		(0.024)		(0.030)
Enrolled in T2		-0.008		-0.043
		(0.028)		(0.031)
Baseline 2016	0.016	0.016	0.060***	0.059***
	(0.018)	(0.018)	(0.016)	(0.016)
Intercept	0.272***	0.275***	0.556***	0.554***
	(0.055)	(0.056)	(0.030)	(0.030)
Obs.	6781	6781	8134	8134
Covariates	X	X	X	X
T2 - T1	-0.0024	-0.0060	-0.0203	-0.0536
(p-value)	0.8278	0.8085	0.0582	0.0498

Notes: Regressions of practice indices (producers only) on T1 and T2. Standard errors clustered at the village level reported in parentheses.

Appendix H: CERP's Terms of Reference with PSDF

ANNEXURE A

Terms of Reference for Commissioning the Centre for Economic Research Pakistan (CERP) for the Evaluation of Two Skill Schemes of the Punjab Skills Development Fund (PSDF)

Introduction

Through this ToR, the Punjab Skills Development Fund (PSDF) is commissioning the Centre for Economic Research Pakistan (CERP) to complete and conclude the RCT based evaluations of two of its skill schemes: (i) the Big Push for Rural Economy (BPRE) and (ii) Skills for Market – Market Linkages (SFM-ML). These schemes, and associated evaluations, were originally approved and initiated under the Punjab Economic Opportunities Programme (PEOP) and have continued through its successor programme the Skills Development Programme (SDP). DFID has provided a funding contribution to both PEOP and SDP, though procurement of the evaluation service provider under SDP will be undertaken directly by the PSDF. Further details are provided below on the PEOP and the SDP programmes, PSDF, and CERP, as well as further background on both the evaluation activities undertaken to date in support of programme implementation and planned evaluation activities for the future.

Background

The Programmes

Punjab Economic Opportunities Programme (PEOP)

The Department for International Development (DFID) and the Government of Punjab (GoPb) initiated the Punjab Economic Opportunities Programme (PEOP) in 2010, a £55m programme that ended in June 2016. PEOP's objective was to increase and diversify income earning opportunities for the poor and vulnerable, initially in 4 of the top 10 poorest districts in southern Punjab. The programme had two main components for achieving its objective: (i) skills development for the poor and vulnerable and (ii) livestock and dairy development for micro-small farmers. Both DFID and the GoPb contributed £25 million each for the two programme components, while DFID contributed an additional £5 million as Technical Assistance (TA) for research, building implementing partners' capacity, and monitoring and evaluation.

The Skills Development Programme (SDP)

CFO COO CERP User Department HOD

1. Based on the success of the skills component under PEOP[1], both DFID and GoPb are currently implementing a follow on skills programme called the Skills Development Programme (SDP). The SDP is a 5 year programme (August 2016 to June 2021) worth £127.5 million, financed 30% by DFID and 70% by the GoPb. The singular focus of SDP is to strengthen cost effective provision of job/income oriented skills for poor and vulnerable people through the Punjab Skills Development Fund (PSDF).

The Implementing Partner

Punjab Skills Development Fund (PSDF)

2. In order to implement the skills component of PEOP, a skills financing fund was established. The Punjab Skills Development Fund (PSDF) became operational in October 2010. PSDF is a not-for-profit company registered with the Securities and Exchange Commission of Pakistan. It is governed by an independent Board of Directors comprising of private entrepreneurs, policy experts, social activists and ex-officio directors representing the GoPb. DFID has observer status on the Board.

3. Under PEOP, PSDF established itself as an effective skills financing model with transparent management and operational systems. PSDF designs and course corrects skill schemes based on market research and findings derived from monitoring and evaluation activities, contracting out the delivery of skills schemes to private and public skill providers on a competitive basis. This approach has been successful in mobilizing the private sector skills providers market and has helped to 'crowd in' diverse private sector players to skills provision. Over 90% of PSDF's skill schemes are delivered through private providers such as training institutes, academic institutions, NGOs and, increasingly, businesses themselves. PSDF has to date skilled over 160,000 poor and vulnerable people (over 35% women) in diverse vocational trades such as: welders, beauticians, industrial-stitchers, school entrepreneurs, farm managers, plumbers, chefs, and computer operators. Most importantly, the success of PSDF has won it high political support and created political interest for reforming the delivery of skills in the public sector and creating a more conducive skills policy and regulatory environment. The Fund remains the only skills financing fund in the country.

4. The SDP programme builds on the success of PSDF under PEOP, with all of the SDP objectives to be achieved through the PSDF which is implementing the programme. With funding support under the SDP, PSDF will expand across all of the 36 districts of Punjab. PSDF will become a sustainable organization that can continue to work effectively post DFID's exit when the SDP

[1] The livestock and dairy component that was being implemented by the Livestock and Dairy Department of the GoPb was closed by DFID and the GoPb following the recommendation of the second Annual Review of the PEOP in 2012.

Legal CFO COO CERP User Department HOD

ends in mid-2021. PSDF will skill 330,000 poor and vulnerable people (40% women) over that time period. Most importantly, based on PSDF's experiences under PEOP, PSDF has sharpened its focus on ensuring all skills lead to jobs/income oriented skills. Efforts for achieving this objective include: i) engaging businesses as training providers on a cost sharing basis, with the stipulation that the business itself will hire at least 50% of the trained individuals, while the remaining trained people will help meet the skilled workforce needs of other businesses ii) having payments to all non-firm skill providers linked to employment targets iii) developing employment facilitation services for students; and iv) designing more innovative skill schemes that focus on connecting the more vulnerable groups - for example small farmers and rural women - with mainstream commercial players and markets.

The Evaluation

5. When PSDF began operations in late 2010, there was limited data of the scale and quality required for PSDF to design skills schemes relevant for both market and students' needs. Granular information required at the district and sub-district levels was lacking on demand for skills, employers' perspective and the scale and quality of skill provision. Available information, most notably the Labour Force Survey provided limited information at the broader provincial but not the district or the sub-districts levels. Additionally, there were limited robust evaluations of large skill programmes covering a general population, or evaluations that would be relevant given PSDF's context.

6. DFID engaged the Centre for Economic Research Pakistan (CERP)[2] in March, 2011 under an Accountable Grant Agreement (AG). The objective was to evaluate returns on skills funded by PSDF following the Randomized Control Trials (RCT) methodology and measuring the economic and non-economic impact of the PEOP Programme. Given the learning opportunity, both within Pakistan and internationally, that an evaluation of a skills programme of this scale offers, CERP offered pro-bono services of its principal research leads with DFID covering all out-of-pocket actual expenses associated with conducting the baselines and the evaluations.

7. In order to fill the information gap for PSDF, and create a baseline for skills evaluation, CERP undertook a number of data gathering initiatives, including an extensive household survey covering 32,000 households in the 4 pilot districts[3] of PEOP, in-depth surveys, census' in 149 villages and representative surveys of 6,000 employers and 3,500 livestock and dairy suppliers.

   Findings derived from these activities were critical in filling information gaps for PSDF and providing a baseline for measuring the impact of PEOP. Most importantly, the findings also identified a significant issue with regards to low uptake of training, particularly by women and especially those from poor and vulnerable households. These findings raised concerns that rural

---

[3] Muzaffargarh, Bhawalpur, Bahawalnagar, and Lodhran. Initially PEOP started operations in these districts.

ABS & CO. Advocates & Corporate Counselors 9-Fane Road, Civil Lines, Lahore

CFO COO CERP User Department HOD

women, a beneficiary group of importance for PSDF, was inadequately participating in PSDF supported training. Low uptake also implied that unless uptake issues were addressed, returns on skills and the programme/PEOP impact could not be measured meaningfully- programme *impact* being the product of *skills uptake into return on skills*

8. The issue of low uptake of skills changed the course of the planned evaluation from being a traditional evaluation of returns on skills offered by PSDF and measurement of the Programme/PEOP impact, to an organic research-evaluation that needed to respond to and be shaped by findings on the ground. This approach has enabled PSDF to design skills schemes that address identified constraints and issues. DFID and the PSDF Board realized this would be a more time consuming approach, but decided to support it in order to address the uptake issues and ensure the evaluation informed implementation by PSDF on an ongoing basis.

9. The subsequent evaluation and experiments revealed that uptake of skills among men could be improved through the provision of employment facilitation services along with skills. However, for women uptake of skills was linked with mobility and social constraints. As vulnerable groups are an important target group for DFID and PSDF, the PSDF Board in its 17th Board meeting (Feb 7, 2014) decided to prioritise the RCT evaluation of the rural women-focused skills scheme of PSDF (and hence focus on Skills for Market[4]) and approved the RCT evaluation of an agricultural skills scheme called Big Push for Rural Economy (BPRE). The BPRE scheme aims to increase the incomes of the micro-small farmers by building their farm, dairy and livestock management skills and linking them with commercial markets.

10. As the focus of the RCT evaluations had shifted to evaluating specific PSDF skill schemes, it became clear that an alternative evaluation mechanism was required to evaluate the overall impact of PEOP. Therefore, in addition to the RCT evaluation of the skill schemes of PSDF, overall PEOP/programme impact continued to be part of the evaluation and was facilitated through a PSDF's GraduatesTracer Study.

11. By the time PEOP closed in June 2016, CERP had conducted the following evaluation related activities:

    11.1. Two follow-on surveys of the SFM-ML scheme. The SFM-ML comprised of the SFM scheme and a market linkage component that was attached to the SFM scheme, in response to trained women's demand for linking them with retailers to whom they could sell their finished garments.

    11.2. Designed two additional evaluations. One focused on determining the most effective ways for PSDF to facilitate employment of the trained (given that CERP had found that the uptake of skills was low among men due to uncertainty regarding job

[4] While all PSDF skill schemes are open for women, SFM has been designed specifically in response to the findings of a household survey which revealed that most women have primary education or less. These women are excluded from formal training, which requires minimal levels of education and does not cater to the socio-cultural contexts of rural women. Initially SFM covered trades such as embroidery, stitching, and food processing. However SFM 2013-14 focused only on stitching given the high demand and completion rates for stitching courses among women as compared to courses in other trades.

CFO COO CERP User Department HOD

prospects post training). The second evaluation was of an agriculture/livestock related skills scheme of PSDF called Big Push for Rural Economy (BPRE).

11.3. Both the evaluation design reports and an Interim Report covering the SFM-ML evaluation and the interim findings have been quality assured by SEQAS.

12. DFID and the PSDF Board agreed that CERP's work required to be closed under PEOP and reinitiated under the new SDP programme in order to complete the RCT evaluations initiated and agreed under PEOP. Hence, the GoPb's Project Document[5] - agreed with DFID for the Skills Development Programme - explicitly mentioned engaging CERP for RCT evaluations under the SDP. There were a number of reasons for engaging CERP again. Substantial time and resources would be lost if a new organization had to be brought on board to complete what was designed and initiated by CERP, CERP possessed the most appropriate experience and expertise in the conduct of RCT based evaluations, as well as the required ground presence and understanding of the local context. It should be noted that PSDF's procurement policy allows single sourcing, on the basis of sound justifications in areas including value for money, transparency and accountability of the supplier.

Purpose, Objectives and Scope

As stated in the summary above, this ToR sets out the requirements of the Punjab Skills Development Fund (PSDF) in commissioning the Centre for Economic Research Pakistan (CERP) to complete and conclude the RCT based evaluations of two of its skill schemes: (i) the Big Push for Rural Economy (BPRE) and (ii) Skills for Market – Market Linkages (SFM-ML). Through evaluation of the BPRE, PSDF as the primary recipient of evaluation findings wants to learn the effective, efficient, sustainable and impactful mechanisms for skilling farmers and what impact a bundle of skills for farmers has on the Village level GDP. Evaluation of the SFM-ML will enable PSDF to learn the most effective, efficient, and sustainable mechanisms for increasing uptake of skills by rural women and the market linkage model that PSDF is implementing for linking skilled women with commercial markets. In addition to economic outcomes, the evaluation will also explore cross-cutting themes and determine the impact of the skills on non-economic outcomes such as: women's empowerment, sense of well-being, state engagement, and aspirations.

13. The evaluations will support PSDF's ongoing monitoring of its skills schemes, enabling it to course correct its skills interventions and improve the composition and design of its skills portfolio. Additionally, the evidence generated will inform the global debate on interventions for farmers and rural women in developing countries. All evaluation products will be delivered on a timeframe agreed by DFID and PSDF on an ongoing basis between now and April 2019, which will ensure they are available to inform programme implementation.

14. This ToR does not cover evaluation of PSDF employment facilitation services. While of interest, more market experience is required before deciding whether to commission an

[5] A Project Document is the Government's equivalent of DFID's Business Case.

ABS & CO. Advocates & Corporate Counsels Branch Road, Lahore

CFO COO CERP User Department HOD

evaluation of this area. Should PSDF decide to proceed with such an evaluation in the future, this would require to be approved by the Technical Assistance Committee (TAC), with the ToR being amended along with a revision of CERP's budget. Further detail on the TAC is provided below.

15. **Independence of the evaluations** will be ensured through a Technical Assistance Committee (TAC) of the PSDF Board. DFID will be represented at the TAC by an Evaluation Advisor, Lead Advisor for SDP and a Programme Manager. All decisions related to evaluations will be taken by the TAC with technical support of the Evaluation Advisor. CERP will share all of its evaluation reports that it submits to PSDF simultaneously with DFID. There are two types of deliverables that will be submitted by CERP during the course of SDP, i.e. i). Final Impact Evaluation Reports and ii). Field Activity Reports including Field Compliance Reports and Survey Field Reports. All final evaluation reports will be quality assured by the Evaluation Advisor. In addition to the Evaluation Advisor, DFID's independent third-party evaluation quality assurance mechanism will quality assure all key evaluation products, i.e. this ToR, design documents[6] and the final evaluation reports of CERP covering the findings and recommendations of the evaluations.

Evaluation of Big Push for the Rural Economy (BPRE)

Background

16. Enabling agriculture and livestock sectors to increase their productivity is an important policy objective of the Government of Punjab's (GoPb) Growth Strategy 2018. The Strategy recognizes stagnant agriculture and livestock productivity as a major development challenge and a determinant of poverty, especially in Punjab's high poverty southern and western districts. PSDF has designed the Big Push for Rural Economy (BPRE) skill scheme to address the challenge of stagnant agriculture and livestock productivity in the high poverty districts of Bahawalpur, Bahawalnagar, Lodhran and Muzaffargarh. BPRE aims to augment productivity through saturation and diffusion of best-practice skills within village-level livestock and agriculture value-chains by directly engaging the leading private-sector companies as training providers working in these sectors in Pakistan. The BPRE scheme also includes a complementary market linkage scheme whereby the corporate skill providers will commit to buying produce from trained farmers. For BPRE's Theory of Change see Annex A1, for the evaluation design of BPRE see the BPRE design document attached with this ToR.

Evaluation Questions

17. Based on the OECD DAC evaluation criteria of impact, effectiveness, efficiency and sustainability, the key questions this evaluation will address are given below. For the purposes of this evaluation these are to be considered the final set of evaluation questions to which no

[6] The design and compliance Reports covering the SPM-ML and the design report for the BPRE have been quality assured by SEQUAS.

CFO COO CERP User Department HOD

further additions will be made, (with the understanding that given results from the field, there will be some flexibility in the extent to which all these questions are addressed).

Suggested questions on Impact:

- To what extent does the agriculture and livestock trainings impact economic outcomes such as: profitability, productivity, consumption, and savings of the individual farmers?
- To what extent does the agriculture and livestock training impact non-economic outcomes such as: social attitudes, empowerment, aspirations, information networks, adoption of skills, and views towards the state?
- What is the incremental impact of access to market on the economic and non-economic outcomes? (To the extent that it is feasible, the evaluation will cover the BPRE market linkage component)

Suggested questions on Effectiveness:

- To what extent would BPRE's overall goals be achieved based on scaling up BPRE if the impact observed in the treatment sample is extrapolated to the broader population at different levels of treatment saturation?

- Do heterogeneous treatment effects between villages in terms of standard welfare measures (e.g. household consumption, employment, children's educational attainment, etc.) suggest factors which determined the extent to which the agriculture and livestock training was able to achieve the objectives of augmenting productivity through saturation and diffusion of best-practice skills within village-level livestock and agriculture value-chains?

Suggested questions on Efficiency:

- What is the cost benefit ratio of the combined agriculture and livestock trainings compared to changes in household consumption and expenditure?
- Are there significant design lessons to be learned from program rollout? Specifically, was the implementation of the programme impacted by external factors not in the programme's control? If so were these risks factors that in the future could be addressed in design?

In order to address the evaluation questions covering efficiency and PSDF will provide CERP with the cost related information for the intervention.

Suggested questions on Sustainability:

ABS & CO. Advocates & Corporate Counsels 9-Fane Road, Lahore

CEO COO CERP User Department HOD

- Were techniques taught in the agriculture and livestock training employed 6 after the training had been completed?
- Do heterogeneous treatment effects in terms of continuation of practices (i.e. there may be some villages where newly-taught practices are sustained and others where they are not used after training) help us to identify factors that determined the sustainability of the trained practices after training was complete?

Activities and Outputs

18. The evaluation comprises of 3 types of activities; (i) compliance activities, (ii) a tracking survey and (iii) preparation of reports. Compliance activities and the tracking survey will be undertaken in a total of 90 villages (60 treatment and 30 control villages).
19. **Compliance Activities (CA)** These activities are the CERP evaluation protocols that the training service providers need to comply with in order to ensure that the evaluation is executed as planned. CA will cover 8,500 treatment households in 60 treatment villages, out of which approximately 6,000 are voucher holders and 2,500 are non-voucher holders. CA will be conducted at the mobilization, enrolment and training stages and will involve, for example, digitization of trainee data, spot-checking during the mobilization process and checking application data to track voucher-holder and non-voucher holders' enrollment. CA are expected to be completed by June, 2018.

20. **Post-Treatment Tracking Survey.** One tracking survey will identify the immediate impact of training on knowledge, practices, output and income for 80% of the sample households in treatment and control villages per tracking survey. An additional end-line survey will measure the medium-term impact on economic and non-economic outcomes on 12,700 treatment and control sample households. These surveys are expected to be completed by November, 2018.

BPRE Outputs/Deliverables to be submitted to PSDF and DFID

21. The following deliverables will be provided in clear and concise English in Ariel font 12. All outputs will be delivered by April, 2019.

Outputs	Due Dates
1. Wheat Compliance Field Report	March 2018
2. Cotton Compliance Field Report	April 2018
3. Livestock Compliance Field Report (all phases)	June 2018
4. Pre Treatment Tracker Field Report	June 2018
5. BPRE Endline Field Report	November 2018

CFO COO CERP User Department HOD

6. Impact Evaluation Report covering the impact of BPRE at the farmer and village level along with policy and programme recommendations.	April 2019

22. The DFID Pakistan Evaluation Advisor will provide quality assurance support for the Compliance and Field Reports. In addition, the DFID Pakistan Evaluation Advisor will also provide quality assurance support as well as procure additional quality assurance via DFID's independent quality assurance mechanism for the final evaluation reports for both the BPRE and SFM-ML

BPRE Evaluation Risks

- Inter-Village Spillover effects: While spillover effects within villages represent a core focus of the evaluation, spillover effects across villages pose a risk to the evaluation, as control villages (which are not provided with the trainings) might be benefitting from trainings conducted in neighboring treatment villages.

- Interference from Third Party Interventions: It is possible that third parties choose to hold interventions while the BPRE intervention is still underway. This is especially problematic if the external intervention is held primarily in control circles, thus breaking the assumption that treatment and control villages, on average differ only in treatment assignment, and consequently dampening our impact estimates.

Evaluation of Skills For Market – Market Linkages (SFM-ML)

Background

23. Augmenting women's earnings, skills and labour force participation is an important objective of the Government of Punjab's Growth Strategy 2018. The Strategy recognizes low female labor force participation and the dependence of rural women on low-wage seasonal activities as major development challenges, especially in Punjab's high poverty districts. Women are an important target group for DFID and PSDF.

24. The Skills for Market scheme has been designed to augment the skills and labour force participation of rural women with low literacy in high poverty districts of Southern Punjab (including Bahawalpur, Bahawalnagar and Muzaffargarh). The SFM-ML scheme has two components: (i) training in tailoring and stitching and (ii) providing market linkages through sales agents. The SFM-ML scheme was complemented with additional services (for example village-based training, and transport) to ease women's social and access constraints. The implementation of the scheme was completed and its evaluation initiated under PEOP.

CFO COO CERP User Department HOD

25. Please see Annex 1B for SFM-ML Theory of Change and the attached SFM-ML design document for more details on SFM-ML evaluation.

Evaluation Questions for SFM-ML Intervention

26. Based on the OECD DAC evaluation criteria of impact, effectiveness, efficiency and sustainability, the key questions this evaluation will address are given below. For the purposes of this evaluation this is to be considered the final set of evaluation questions to which no further additions will be made, with the understanding that given results from the field, there will be some flexibility in the extent to which all these questions are addressed.

Suggested questions on Impact:

- To what extent does the SFM-ML skills training and market linkage component impact economic outcomes such as: individual earnings, household consumption, and income?
- Non-economic outcomes such as: state engagement, sense of well-being, empowerment and aspirations?

Suggested questions on Effectiveness:

- What were the cost-benefit ratios of per-trainee cost for SFM-ML skills training and the market linkage component compared to changes in economic outcomes such as: individual earnings, household consumption, and income?
- Do heterogeneous treatment effects, i.e. differences in how well the training works in different regions or for different demographic groups, suggest key factors which influence how well skills training and market linkages work at (a) augmenting the skills and (b) increasing labour force participation of rural women with low literacy in high poverty districts of Southern Punjab.

Suggested questions on Efficiency:

- What is the ratio of per-trainee cost of SFM-ML skills training and market linkage compared to the per-trainee changes in income?
- Are there significant design lessons to be learned from program rollout? Specifically, was the implementation of the programme impacted by external factors not in the programme's control? If so were these risks factors that in the future could be addressed in design?

In order to address the evaluation questions covering efficiency, PSDF will provide CERP the required cost per trainee information.

Legal | CFO | COO | CERP | User Department HOD

Suggested questions on Sustainability:

- What share of those trained in the SFM-ML skills training and market linkage treatment were still engaged in stitching for the market 6 months after the training had been completed? How much did market linkages affect this number?
- Do heterogeneous treatment effects suggest key factors which determine the sustainability of the impact of the training after its completion. There may be some individuals who continued to use the newly-learned skills and others who did not. The difference between them can tell us about which kinds of individuals in which kinds of places are likely to benefit from the training for the long term.

Activities

27. As mentioned earlier in this ToR, the evaluation of the SFM-ML was initiated under PEOP. The SFM and SFM-ML Evaluation Design and Compliance Activities (CA) reports were quality assured by SEQAS and are attached with this ToR. To date, 1 survey and 3 tracers have been conducted for the SFM-ML evaluation. These include a baseline survey (October-December 2013) and post training first and second follow up tracers (December 2014-January 2015 and November-December 2015 respectively). The surveys were funded under PEOP. A report covering the interim results of the evaluation has been quality assured by SEQAS. The end-line survey covering 10,700 households will be conducted under SDP and the final report is expected to be submitted to the PSDF and DFID for quality assurance by April 2018. The final report will provide more conclusive findings covering the sustainability, efficiency, effectiveness and impact of the scheme along with programme and policy recommendations.

Outputs/Deliverables

28. All deliverables will be in clear and concise English and in Ariel Font 12. The following outputs will be submitted to PSDF and DFID no later than April 30, 2018.

Outputs	Due Date
1. Field Report on the SFM-ML End line survey	Both reports will be delivered no later than April 30, 2018
29. Final Report on the impact of SFM-ML on women covering programme and policy recommendations related to efficiency, effectiveness, sustainability and impact of the different experiment arms for increasing uptake of skills for women and their market linkages.	

ABS & CO. Advocates & Corporate Consultants Legal 9-Fane Road, Lahore

CFO COO CERP User Department HOD

30. The DFID Pakistan Evaluation Advisor will provide quality assurance support for the Compliance and Field Reports. In addition, the DFID Pakistan Evaluation Advisor will also provide quality assurance support as well as procure additional quality assurance via DFID's independent quality assurance mechanism for the final evaluation reports for both the BPRE and SFM-ML

Risks

31. Evaluation specific risk for Skills for Market – Market Linkage is mainly related to the inability of the middlemen/women to sustain profits after the subsidies have been removed, as a result of which they may stop working with the ML seamstresses.

Data Availability

32. Anonymized data from all surveys, tracers and on treatments will be made publicly available on CERP's website as provided in the partnership agreement.

Risks - Overall

33. Overall risks covering both of the evaluations are as follows:
    - Security risks of the personnel engaged in field work.
    - Country Risk: Untoward events may impact the roll-out of the project and hence delay or adversely affect the viability of the evaluation.
    - Training Service Provider Compliance Risk: A potential risk is implementation compliance by training service providers (TSPs). The evaluation design protocol is complicated and requires the training service providers to check their usual processes in a number of ways.
    - Policy Uptake risk: The findings may have little policy impact. This risk is largely addressed by the strong ownership of this exercise by PSDF and their commitment to learn from the findings of the evaluations and inform the development of their program by this evidence.

Audience

34. The evaluations are intended to benefit the future design efforts of PSDF, Government of Punjab and DFID. Other recipients include policy makers in Federal and the other Provincial Governments and development partners. The evaluation findings will also be of interest to the wider global development community working on skills development.

Further Deliverables

CFO COO CERP User Department HOD

- **Synthesis Report:** The Synthesis Report will summarize the findings and recommendations of both of the RCT- based evaluations covered in this ToR. This report will be submitted for quality assurance by end of May 2019.
- **Two policy briefs** using results from impact evaluation reports and syntheses reports

35. In addition to the reports submitted for quality assurance the evaluation will lead to several academic publications, papers and policy notes that serve to transfer knowledge developed in this setting to the broader development community. DFID and PSDF will have unlimited access to the material produced by the supplier (as expressed in DFID's general conditions of contract).

Engagement and Communications

36. Regular meetings between CERP and PSDF will take place during the course of implementation and evaluation to ensure that both parties are on board with the design and implementation procedures and protocols. The final impact evaluation results will be presented in addition to PSDF to the Technical Assistance Committee (TAC) of the PSDF Board. TAC will be kept informed of all progress and field challenges and issues requiring technical advice and approvals.

37. Findings will be disseminated by CERP through local and international seminars, conferences and policy dialogues. Policy dialogues will be organized to enable policy makers understand the policy implications of the findings and the recommendations and promote partnerships for effective policy solutions. In addition, CERP's policy reports and briefs will also be disseminated and document the results. A special effort will be made to share results with relevant policymakers and stakeholders from other provinces in Pakistan.

X. Governance, reporting and contracting arrangements

38. For the coordination of its field work, CERP will set up communication protocols with PSDF's management. CERP will report to the TAC of the PSDF Board: i) Evaluation results of the final SFM report ii) Overview of compliance activities in July 2018 and then the final evaluation results of BPRE in April, 2019. In addition, TAC will review and approve any changes to this ToR or to the design and implementation of the evaluations.

39. PSDF will ensure that the Training/Service Providers engaged as part of this work provide CERP access to relevant data and information including per trainee costs. PSDF will help build networks between CERP and other relevant actors to ensure support, complementarity, and improved coordination.

CFO COO CERP User Department HOD

XI. Skills and Qualifications

40. CERP is expected to set up an evaluation team comprising of national and international experts with the following expertise and experience:

- Sound understanding of RCT evaluation design and methods and accurate interpretation of data. The team will primarily require competence in quantitative methods, in particular the design and implementation of RCT evaluations, supplemented by data analysis and potentially qualitative skills to supplement core quantitative findings.
- The team will possess excellent written skills, and have a demonstrated ability to communicate complex studies and findings in an accessible way for non-technical readers, including presentation of data in visually appealing ways, highly structured and rigorous summaries of research findings and robust and accessible synthesis of key lessons from across different studies. Flexible approach to ensure that the evaluation design is closely linked to the skills schemes' design and their implementation by PSDF.
- Strong presence in and experience of field-work in South Punjab.
- Expertise and understanding of skills sector reforms and ability to assist governments to create policy using feedback loops by applying impact evaluation to program design and calibration.

XII. Quality Assurance

- All CERP deliverables will be quality assured and signed off by CERP's principal investigators before being submitted to PSDF and DFID.
- All deliverables will be submitted to DFID and PSDF simultaneously by CERP. PSDF will provide its feedback on the deliverables to DFID within 10 days of receiving them.
- Key CERP deliverables specified in this ToR will be quality assured by DFID's Evaluation Advisor. In addition, final evaluation reports will also be quality assured by DFID's third party evaluation quality assurance mechanism i.e. Evaluation Quality Assurance and Learning Service (EQUALS). These include: The impact evaluation report of the SFM-ML, the impact evaluation report of the BPRE, and the synthesis report.
- DFID's Evaluation Advisor will keep the TAC informed of the quality of the outputs received from CERP.

Payment of Fees

ABS & CO. Advocates & Corporate Counselors 9-Faisal Road, Lahore

CFO COO CERP User Department HOD

- Payment will be made by PSDF to CERP as per payment Schedule provided in the Partnership Agreement.
- CERP will produce the deliverables on time unless under exceptional circumstances that are communicated to PSDF and DFID (as early as the likelihood of delay become apparent to CERP) and following new timelines that are agreed by PSDF and DFID.

XIV. Budget

Referring the Clause 7C(iii) ***"Monitoring, Research, Evidence & Evaluation"*** of the Project Document for SDP, an amount of up to £1.0 million has been allocated for evaluations. However, PSDF will ensure realistic market costs estimates and value for money in negotiating the price of the deliverables and the overall budget with CERP.

XV. Duty of Care

- CERP will be responsible for managing all logistics, including any in-country travel arrangements for the purposes of this evaluation. It will be responsible for the duty of care, safety and well-being of their personnel.

XVI. Key Documents

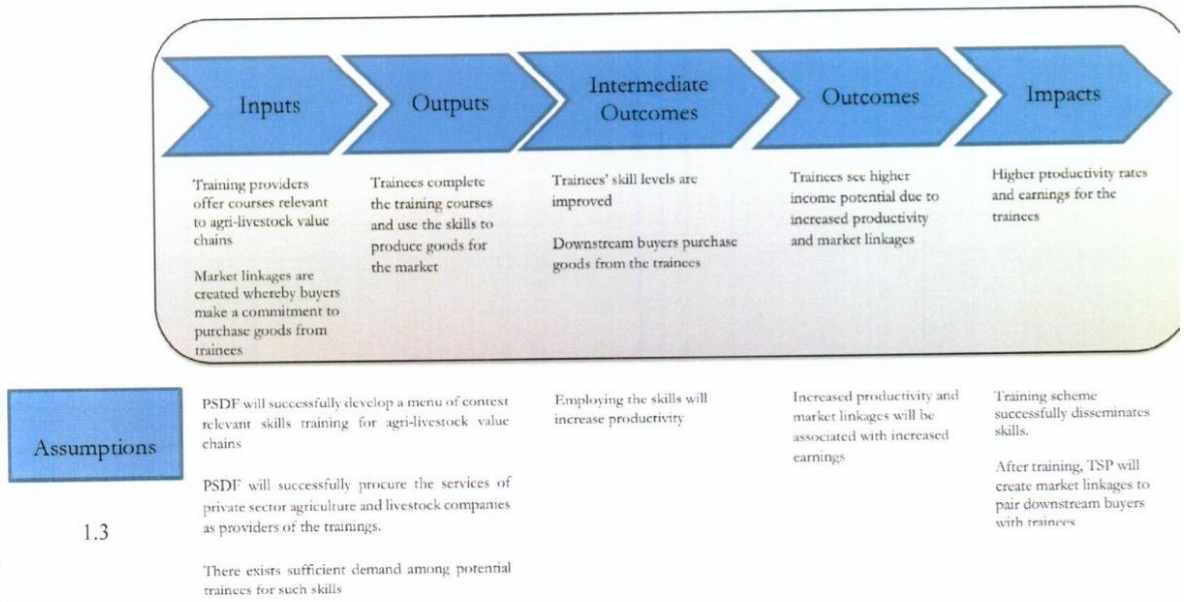
- Following is the list of all key documents, which are provided with the TORs:
- Closure Report of CERP (following closure of PEOP)
- BPRE Evaluation Design Report
- SFM 2013-14 Design and Compliance Reports
- SFM – Market Linkage Design and Compliance Reports
- SFM-ML 2013-14 Interim Impact Evaluation Report

AABS & CO. Advocates & Corporate Consultants 94 The Mall Road, Lahore

CFO COO CERP User Department HOD

1.2 Annexure 1.A

Figure 1: BPRE and Market Linkage Theory of Change



1.3

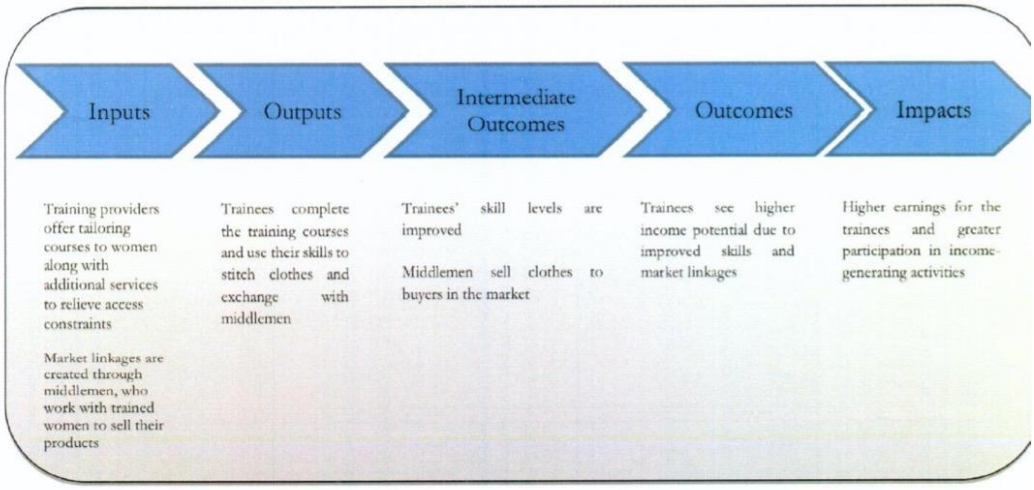
ABS & CO. 9-Fane Road, Lahore Advocates & Corporate Counsellors

32

CFO COO CERP User Department HOD

1.4 Annexure 1.B

Figure 2: Skills for Market and Market Linkage Theory of Change



ABS & CO. Advocates & Corporate Counselors 9-Fane Road, Lahore

Legal CFO COO CERP User Department HOD