PUNJAB ECONOMIC OPPORTUNITIES PROGRAM

Interim Report on Impact Evaluation of SFM 2013-14

Ali Cheema
Lahore University of Management Sciences

Asim I. Khwaja
Harvard University

Muhammad Farooq Naseer
Lahore University of Management Sciences

Jacob N. Shapiro
Princeton University

CERP | Center for Economic Research in Pakistan

Contents

Table of Figures

Table of Tables

Acronyms

BPRE	Big Push for the Rural Economy
CERP	Center of Economic Research in Pakistan
DfID	Department for International Development
GoPb	Government of Punjab
GT	Group Transport
ITT	Intent to Treat
IV	**Instrumental Variables**
LATE	Local Average Treatment Effect
ML	Market Linkage
NGO	Non-Governmental Organization
PEOP	Punjab Economic Opportunities Program
PKR	Pakistani Rupee
PML-N	Pakistan Muslim League Nawaz
PPP	Pakistan People's Party
PCR	Program Completion Review
PTI	Pakistan Tehrik-e-Insaf
PSDF	Punjab Skills Development Fund
RCT	Randomized Control Trial
SD	Standard Deviation
SFE	Skills for Employability
SFM	Skills for Markets
TSP	Training Service Provider
VBT	Village Based Training
UC	Union Council

Executive Summary

This report outlines interim findings from an impact evaluation of the Skills for Market (SFM) scheme carried out in 2013-14 in three high-poverty districts of Southern Punjab.[1] SFM 2013-14 specifically targeted rural women and aimed to improve their livelihood and labor force participation through need-based skills training. SFM 2013-14 sought to alleviate access constraints faced by rural women and to identify ways to do so at the lowest possible cost. The evaluation sample included 9,893 households in 324 villages.

An earlier evaluation of SFM 2013-2014 tested the notion that travel constraints are a major barrier to skills acquisition, especially for rural women and measured the impact of different methods of removing such constraints and delivering training on participation rates. That report, "Alleviating Take-up Constraints for Rural Women," (Cheema et al. 2015) highlighted four major findings. First, it is extremely challenging for rural women to travel for vocational training. Second, many of those challenges cannot be monetized; providing stipends to help cover the fiscal and social costs of training increases participation but only modestly. Third, providing in-village training significantly increasing uptake amongst women in rural areas. Fourth the fact that training suppliers located their training centers in randomly specified villages for that impact evaluation proves it is feasible to scale accessible training for rural women.

This interim report is based on the impact evaluation of SFM 2013-14, which is a component of PEOP. The results demonstrate that making skills training accessible for rural women increases uptake and that such training can successfully provide valuable income-generating skills at the individual report. In particular, this report focuses on the impact of the training on three outcomes as measured through a randomized control trial (RCT) methodology:

1. Trainee skills acquisition as measured through engagement in the vocational skill;
2. Downstream socio-economic outcomes including earnings, employability, civic engagement, physical and mental health, gender equality and government services usage
3. A revealed preference measure of satisfaction with the program, voting behavior.

Our analysis shows that SFM 2013-2014 trainings resulted in a marked increase in stitching activity; trainees were significantly more likely to engage in stitching for pay and were more likely to have stitched at all. We do not detect impacts on non-stitching downstream outcomes to date using surveys. We do find evidence of satisfaction with the program using a revealed preference approach, where voting outcomes is seen as a proxy for citizen satisfaction[2]. Specifically, results showed that voters in villages with village-based training were likelier to vote for the ruling political party, suggesting citizen satisfaction with the scheme.

[1] Bahawalpur, Bahawalnagar and Muzaffargarh.

[2] There are a number of issues with using subjective satisfaction with the program as an outcome variable. There is no program in control circles nor at baseline, so it is extremely artificial to ask survey respondents to rate their satisfaction with a program they have not experienced. Moreover, it is not clear that self-expressed satisfaction would be behaviorally relevant.

From the initial evaluation, we suggest making training more accessible for rural women by placing training centers in rural villages, which will lead to an increase in the uptake. The results also suggest that such training schemes are successful in increasing stitching activity and skills level of the participants. Given that this is an interim report we refrain from an extended discussion of recommendations and policy implications

The report will be updated with additional information on household-level impacts once a full endline household survey is complete in Fall 2016. For a training program such as this one many of the most important incomes may come at the household level in terms of intra-household time allocation, education decisions for children, marital outcomes, and the creation of new businesses that leverage skills conferred through SFM 2013-14. Program impacts on those outcomes as well as long-run impacts on household consumption and investment decisions will be measured in the final endline survey.

1 Introduction

Skills for Market (SFM) scheme specifically targets rural women to improve their livelihood and labor force participation through need-based skills training, with an aim to test the hypothesis that access and travel constraints are a major barrier to skills acquisition. Initiated in 2012, over 7000 rural women are already trained under the SFM. Given the hypothesis that women face significant barriers to out-of-village travel, the evaluation of SFM 2013-14 is designed to test different ways of alleviating those constraints.

This report outlines interim findings from an RCT based impact evaluation of the Skills for Market (SFM) scheme carried out in 2013-14. We study the impact of SFM 2013-14 skills based training on skill acquisition, downstream outcomes including earnings, employability, civic engagement, physical and mental health and government services usage, and the general satisfaction from the scheme, measured through a revealed preference approach in the form of voting behavior.

This report is structured as follows: Section 2 provides the broader context of the SFM in terms of its history and players involved; Section 3 details the evaluation methods; Section 4 describes the implementation of SFM 2013-14 scheme; Section 5 provides the broader project timeline; Section 6 presents the findings from evaluation whereas Section 7 concludes.

2 Context

This section describes SFM 2013-14, its history and its objectives in detail. It begins by briefly summarizing Punjab Economic Opportunities Program (PEOP) and the Punjab Skills Development Fund (PSDF) in Section 2.1, followed by describing collaboration between CERP, PSDF and DfID in Section 2.2. Section 2.3 background of the evaluation and gives evidence from field visits and prior evaluations that informed the design of SFM 2013-14 scheme and the objectives of evaluation; and Section 2.4 presents the theory of change.

2.1 Background on Punjab Economic Opportunities Program

In 2010 the Government of Punjab (GoPb) in collaboration with the Department for International Development (DfID) initiated the Punjab Economic Opportunities Program (PEOP) in four of the poor districts of South Punjab[3] (also referred to as 'pilot' districts) to help improve incomes of the poor and vulnerable and reduce poverty and vulnerability. In order to achieve these objective, PEOP had two program components i.e. i) increasing the access and returns to livestock through Livestock and dairy Development (L&DD)[4] and ii) increasing employability and earnings of low income, poor and vulnerable families by augmenting technical and vocational skills. In order to ensure effective implementation of the two components, DfID also committed to providing Technical Assistance (TA). DfID has provided £25 million for training and £5.0 million as TA. The GoPb has matched DfID's funding for training equivalent in Pak Rupees. The total value of PEOP is £55.0 million.

[3] Bahawalpur, Bahawalnagar, Lodhran and Muzaffargarh

[4] The L&DD component was closed down following the recommendations of the second Annual Review of the PEOP.

In 2012, following the recommendation of the second Annual Review of PEOP, DFID and the GoPb decided to close the L&DD component of PEOP. Funds marked for this component were diverted to the skills component of the PEOP. Hence, after the closure of the L&DD component, PEOP comprised of only the skills component and the TA from DFID which aimed to assist the implementation, monitoring and evaluation of the skills component.

To implement the skills component of PEOP, a skills financing fund was established to implement the skills component. The Punjab Skills Development Fund (PSDF) had the target of skilling 145,000 poor and vulnerable people (40% women) by June 2016 in trades that will lead to employment and improved incomes. PSDF was created to increase the access of low income, poor and vulnerable members of the society to vocational training and skills acquisition programs with an aim to achieve the following outcomes at the household level:

- Increase income earning potential
- Increase access to employment opportunities and employability
- Increase participation of women and other marginalized groups in the labor market

PSDF performs core functions in-house and ancillary functions have been contracted out to independent third parties. For example, PSDF designs training schemes based on market research and evidence, it prices the skills schemes, and contracts them out on a competitive basis to skill providers. Market research, monitoring of the training providers and evaluation of the skill schemes has been contracted out to independent third parties.

PSDF is a not-for-profit company established under the Companies Ordinance 1984 and registered with the Securities and Exchange Commission of Pakistan. PSDF has an independent Board and management. The Board comprises of private entrepreneurs, social activists, policy experts, politicians and ex-officio government representatives. Since its establishment, PSDF has expanded from its initial 4 districts to 10 additional districts in central and northern Punjab (as referred to as 'expansion' districts). PSDF has established itself as an effective business model for a skills financing fund. The Fund has mobilized private sector training providers through its competitive contract awarding mechanisms and is on track to achieve its training target. Most importantly, PSDF enjoys high political support and has raised the political appetite for wider skill sector reforms.

The principal skills schemes of PSDF are Skills for Jobs (SFJ) and Skills for Markets (SFM). SFJ offers certified training in trades with demand in mainstream commercial markets. Training requires a minimum level of education and takes place in the formal classroom environment. SFJ is aimed predominantly at the relatively educated poor and vulnerable people – both men and women, though the majority of the students have been men. In order to target the more marginalized women with limited or no education at all, PSDF designed the SFM scheme. Training is provided in an informal setting and does not require any minimum qualification. Trades covered include predominantly ones of interest to women such as tailoring, embroidery, and clothes embellishment.

The GoPb and DfID's current funding commitment to PSDF will come to an end on June 30, 2016. However, both the GoPb and DfID will continue to support the skills sector in Punjab, including training through PSDF, through a follow on skills program for another 5 years (2016-2021) through a follow-on skills program called Skills Development Program (SDP). Under this support, PSDF will be established as a sustainable skills financing fund with operations across all of the 36 districts of Punjab.

2.2 Collaboration with the Center for Economic Research in Pakistan

The Government of Punjab, DfID and PSDF entered into a collaborative arrangement with Center for Economic Research in Pakistan (CERP) to conduct independent evaluations of PSDF's skills schemes, as well as to provide evidence-based input to PSDF that can be used to help PSDF design skill schemes.

The Center for Economic Research in Pakistan (CERP) under an Accountable Grant Agreement with DfID is conducting independent third party rigorous scientific impact evaluation of PSDF's portfolio of skill schemes with an aim to support effective program calibration.

It is important to note that CERP's evaluations of PSDF's schemes uphold ethical considerations, as given below:

1. Ethics Approvals for the research team is obtained from institutions with reputed and well-developed ethics approval systems such as Harvard or Princeton's Institutional Review Boards.
2. Research and evaluation is relevant and high quality with clear developmental and practical value: CERP has taken a number of steps to ensure relevance, quality and practical value. The relevance of the evaluations is ensured because the interventions being evaluated form the core of a large-scale skills development program that is being supported by public funds.
3. The practical value of the evaluations is ensured as it is dealing with core interventions of a major public policy program in a critical developing economy.
4. Harm is avoided to participants in the studies and confidentiality of information, privacy and anonymity of study participants is insured.
5. Strict protocols are put in place that safeguard the privacy and confidentiality of respondents, all personally identifiable information is only saved in encrypted password-protected files.
6. All benefits and services that are made available as part of an evaluation are presented to the prospective trainees transparently, both in writing and verbally. The receipt of this information is recorded as part of the intervention roll out protocol and the data is back-checked and verified on a sample basis.
7. Participation in the research and evaluation is voluntary and free from external pressure: We follow strict protocols to provide all information about the treatments being offered in writing and verbally. During the intervention roll-out, a consent script is read and a

household is only enrolled in the evaluation if consent is given. Verbal consent will avoid complications that arise from working with less literate populations. This activity is conducted with trainees within their household and away from community pressure.

8. Protocols are designed to respect cultural sensitivities: This is ensured by piloting these protocols and obtaining feedback on them through qualitative fieldwork. This is an important part of the treatment compliance activity.
9. In the case of evaluations CERP is responsible for the evaluation design and oversight of the implementation to ensure compliance with the evaluation protocols, which is conducted through CERP field coordinators or third parties that are independent of PSDF and the TSPs.

It is also important to note that CERP's evaluations do not include the evaluation of PSDF as an organization, or an overall evaluation of PSDF's skills schemes. CERP under an Accountable Grant Agreement with DfID is conducting independent third party scientific impact evaluation of PSDF's select skill schemes to support effective calibration of PSDF's skills interventions. The partners recognize that cost effective impact requires interventions that are grounded in and informed by solid evidence and the need to address evidence gaps on both the demand and supply sides of the skills and the labor markets.

This collaboration is recognition of the fact that cost effective impact requires interventions that are grounded in and informed by solid evidence and address issues on both the demand and supply sides of the skills and labor markets. The key components of this collaboration include:

- Producing rigorous evidence to enable PSDF to devise evidence-based and empirically grounded design of an integrated portfolio of skills interventions in the market for labor and skills training; and
- Continuous monitoring and evaluation of the impact of skill schemes of PSDF using rigorous randomized-control-treatment (RCT) methodology to enable recalibration for maximally effective delivery.

The evidence-based component ensures that the design of interventions is informed by and responds to opportunities and constraints embedded in the specific local context in which both markets operate. For this purpose, a large scale baseline survey was conducted to inform program design. This Baseline Phase included village and household surveys, employer surveys and trainer surveys in the program districts. Moreover, pilot interventions of PSDF were carried out that provided detailed insights into the demand and requirements for skills and the constraints faced in skill acquisition and employability.

The Evaluation Phase of the collaboration requires CERP to rigorously evaluate the impact of PSDF's main skill schemes on economic and non-economic returns in the target population of PSDF. It includes evaluation of four distinct skills schemes of PSDF that are listed in Table 1 below along with their target population, focus and current stage.

Table 1: CERP's evaluation of the four PSDF schemes

PSDF's Scheme evaluated by CERP	Target Population of the Scheme	Focus of the Scheme	Stage of the scheme	Expected Deadline of the scheme (including implementation followed by evaluation)
Skills for Market (SFM) 2013-14	Rural women in poor and vulnerable households	Skills training in domestic tailoring	Evaluation	December 2017
Market Linkages 2015	Subsample of SFM 2013-14 graduates	Market access	Evaluation	December 2017
Skills for Jobs (SFJ)	Urban men	Skills training in menu of popularly demanded technical courses	Pre-implementation	TBD
Big Push for Rural Economy (BPRE)	Rural men and women	Skills training across the agri-livestock value chain	Pre-implementation	December 2018

As part of the collaboration, the ownership/copyright of this data and analysis/reports belongs to CERP, while complete access is provided to DfID, as a result of which DfID is allowed to share and publish analysis/reports and limited survey data after consulting with CERP (DfID can publish full survey dataset 2 years after the date of completion of project).

As part of the collaboration with DfID, CERP has submitted design reports for the training schemes mentioned in Table 1. Particularly relevant to SFM 2013-14, CERP has submitted two reports to DfID. "SFM 2013-14: Design and Compliance Report" describes the overall scheme as well as sample characteristics, whereas "Alleviating Access Constraints for Rural Women" analyzes course uptake across treatment arms and highlights which treatment is most successful in alleviating access constraints and increasing course uptake among rural women.

- **Skills for Market 2013-14: Design and Compliance (Cheema et al. 2014):**
  This report gave the broad overview of demographics of the households and trainees in sample villages as well as the description of the various treatments that PSDF introduced in the SFM 2013-14 design. It was agreed with DfID that the design report will serve as the de facto terms of reference for this report.
- **Alleviating Access Constraints for Rural Women (Cheema et al. 2015):**
  This report analyzed the numbers for course uptake across treatment arms, based on the evidence from SFM 2013-14. It showed that:

- Rural women find it difficult to travel to training centers due to logistical hassle and social norms that limit their mobility.
- These challenges cannot be monetized entirely; stipends can only modestly help covering fiscal and social costs of training to boost course participation modest.
- Providing in-village training significantly increases uptake among rural women
- The fact that training suppliers located their training centers in randomly specified villages for that impact evaluation proves it is feasible to scale accessible training for rural women.

2.3 Background

The design and evaluation of SFM 2013-2014 builds upon previous evaluations by CERP of PSDF schemes, described below. For all these schemes, including SFM 2013-14, an Encouragement Design was used where a randomly selected sub-group from the representative survey households in the Program districts was offered a voucher to enroll in PSDF's course. The vouchers guaranteed admission in the course.

2.3.1 Evaluation of Skills for Employability (SFE) 2011-12

In late-2011 PSDF launched the Skills for Employability (SFE) 2011-12 training scheme based on a large expressed demand for skills acquisition from baseline surveys conducted in early-2011. CERP's evaluation of the scheme revealed low uptake from the general population. This was especially serious for females, as only 5% of women offered vouchers for training ended up enrolling in courses. Low uptake was especially acute for women belonging to the poor and vulnerable households, as well as those living further from training centers (Cheema et al. 2012c). These findings raised concerns that the sub-populations of interest to the training scheme were not participating enough to benefit from the trainings supported by PSDF.

2.3.2 Evaluation of Skills for Markets (SFM) 2012-13

After training course under SFE 2011-12 ended, CERP conducted several qualitative follow-up interviews and focus groups to determine the reasons for low uptake for women. This research revealed that the low uptake was not due to the lack of demand for training. Instead, women did not use their vouchers because of a host of social and practical constraints such as community norms against travel, household obligations, and logistical hassles in rural areas.

Based on the focus group findings, PSDF designed two directed pilot interventions that could mitigate access constraints for rural women and thereby improve uptake and tested them in the Skills for Market 2012-2013 training scheme: (1) reducing distance constraints by placing training centers within rural villages village instead of the standard practice of setting up training centers in larger towns; and (2) addressing social constraints by hiring NGO social mobilizers to hold focus groups with women to stress the course's usefulness and encourage them to participate.

The RCT-based SFM 2012-2013 evaluation showed that, enrollment rates increased by 35 percentage point for women who had the training center located inside their villages, and 17 percentage points for women who received social mobilization (17%), but enrollment rates stayed

low for women who received only information (2.6%) (Cheema et al. 2013a). A comparable pattern was witnessed for the course completion rates. Women who acquired training in their own villages had an overall 28% course completion rate followed by women who received social mobilization with a course-completion rate of 17%. In contrast, only 1% of women who only received standard information completed the course.

These numbers show that alleviating access constraints by either decreasing distance or by providing social mobilization helps in enrolling and retaining greater proportion of women in the training scheme. CERP conducted post-training focus groups to understand why women who did not have a training center in their villages had worse uptake rates than women who lived in a village with a training center. The logistical hassle of traveling to the training center was the most common reason cited for low uptake and course completion rates.

This evidence raised practical concerns. Although locating the training center in the village significantly increased uptake, it did so at a higher cost. PSDF's Board recognized the large welfare gains associated with finding solutions that increase women's access to training in a cost-effective manner and demonstrating their logistical viability at a large scale. The next SFM training scheme, SFM 2013-14, was therefore designed to reach more women and to identify cost-effective interventions that could alleviate distance related access constraints for villages where a training center was not located.

2.3.3 Constraints to Accessing Training

Findings from PEOP baseline surveys, field visits and previous evaluations including SFE 2011-12 and SFM 2012-13, provided strong evidence that women in Southern Punjab have low enrollment rates due to access constraints, despite having strong demand for skills trainings. These findings led to additional intensive field visits to inform the design of SFM 2013-14 training scheme. The purpose of the field visits was to:

- Get qualitative feedback on different limitations women face in accessing skills training; and
- Assess the practicality of different solutions aimed at alleviating these constraints.

These field visits were conducted from September to November 2013. The field team tasked with these visits comprised of CERP researchers as well as staff of the local survey firm. For each visit, the field team drew a random sample of 20 villages from a total set of 356 villages[5] and visited 1-2 households in each village, also selected randomly (i.e. approximately 30 households per visit). A randomized approach is used to ensure that information collected is representative of the pilot districts.

During these visits, the field team interviewed two sets of respondents: 1) male and female household members, and 2) influential members of the community. While interviewing the household members, field team asked household members how they would rank a list of possible

[5] These villages consist of all villages in evaluation sample for SFM 2013-14 (324 villages; See Section 3.3.1) and 32 villages in the SFM 2012-14 treatment sample (Cheema et. al 2013a)

solutions to access constraints. After the interview, the field team asked the household members as well as other people in the village to identify key influential members of the community. Field team then conducted interviews with the influentials most frequently named and elicited their preferences as well. In order to respect prevailing gender norms and to encourage honest responses, female field researchers conducted the interviews with female members whereas male field researchers conducted interviews with the male household members and influential members of the community.

In addition to the field visits, CERP research staff also compiled an extensive literature review of 80 scholarly papers from 1990-2013 (CERP 2013) that studied access constraints (distance, household responsibilities, information deficits, cultural norms etc.) faced by women and their solutions in context of developing countries across Africa, Latin America and South Asia (including Pakistan). The literature review was conducted from June to July 2013.

The field visits and literature review identified four main access constraints:

1. Physical Distance: In field visits and previous RCT-based evaluations (SFE 2011-12 and SFM 2012-13), physical distance to the training center emerged as one of the main reasons for lack of enrollment in or completion of the training scheme (Cheema et al. 2012b). The SFM 2012-13 RCT evaluations with rural women found that for every kilometer increase in distance, enrollment fell by 4 percentage points. The evaluation also found that 46% of the trainees who refused to participate in the training scheme stated distance as the primary constraint (Cheema et al. 2013a).
2. Safe and Reliable Transportation: While literature often cites distance to the training center as a major barrier to course participation (Solotaroff et al. 2012; Maitra and Mani 2012; Kabeer et al. 2012), this constraint consists of more than just geographic distance. In the context of rural women, lack of safe and reliable transportation presents an additional logistical challenge. This fact was stressed by the male household members in field visits. These individuals refused to send women to training centers in other villages in the absence of familiar drivers and reliable transport facility.
3. Financial and Credit Constraints: For rural women, participation in the training schemes may imply additional traveling costs or the potential income loss due to the opportunity cost of time allocated to the training. Lack of adequate monetary incentives to compensate for these opportunity costs can also prevent women from participating in the course. This assumed risk, (i.e. opportunity cost of participating) must be compensated in order to increase enrollment rates (Cheema et al. 2014). This constraint was the second most cited reason for course dropout in the SFM 2012-13 training scheme; 20% of the dropouts cited low stipend as the primary reason for dropping out of the scheme (Cheema et al. 2013a).
4. Social Norms: Restrictive social norms also present an additional barrier to access for rural women (Wigfield and Turner 2012). Crucially, men see transgressing restrictive gender norms as impacting their reputation directly (Jamali 2009), and may be unwilling to allow women of their household to participate in skills training or a potential income generating activity, even if they see its value (Naqvi and Shahnaz 2002). To mitigate this constraint, a

pilot was designed as part of the SFM 2012-13 to test the impact of Social Mobilization on enrollment rates for rural women in the PEOP program area using a RCT design. As part of this intervention, households and their female members were mobilized collectively and individually by trained social mobilizers in villages that did and did not receive a training center. This treatment was aimed at reducing the extent of social constraints faced by rural women by mobilizing their household members. The enrollment rates increased by 14% as a result of social mobilization in villages that did not have a training center within them (Cheema et al. 2013a).

SFM 2013-14 is a skills training scheme designed to augment the human capital of marginalized women in South Punjab. Given the hypothesis that rural women face significant barriers while travelling, especially out of village, the primary objective of this evaluation was to analyze the impact of various treatment options aimed at alleviating distance constraints on uptake of training by rural women.

Rural women are a principal focus of this training scheme since more than 80% of the rural women in the pilot districts live in acute poverty and have below-primary levels of education (Cheema et al. 2012a). Problems of low mobility and social barriers preclude these rural women from accessing training. Furthermore, baseline surveys and previous evaluations reveal that in spite of strongly expressed demand, women from the target population have low enrollment rates in training courses due to access constraints.

The objective of conducting an evaluation of the SFM 2013-14 skill scheme was therefore to measure:

a. The impact of different design calibrations on participation – thereby figuring out how to address obstacles that adversely impact uptake in a social context where women face severe social, distance and financial constraints; and
b. The impact of encouraging women to attend popularly demanded training in tailoring on their economic and non-economic well-being.

The objectives of evaluation will be explained further in Section 3.

2.4 Theory of Change

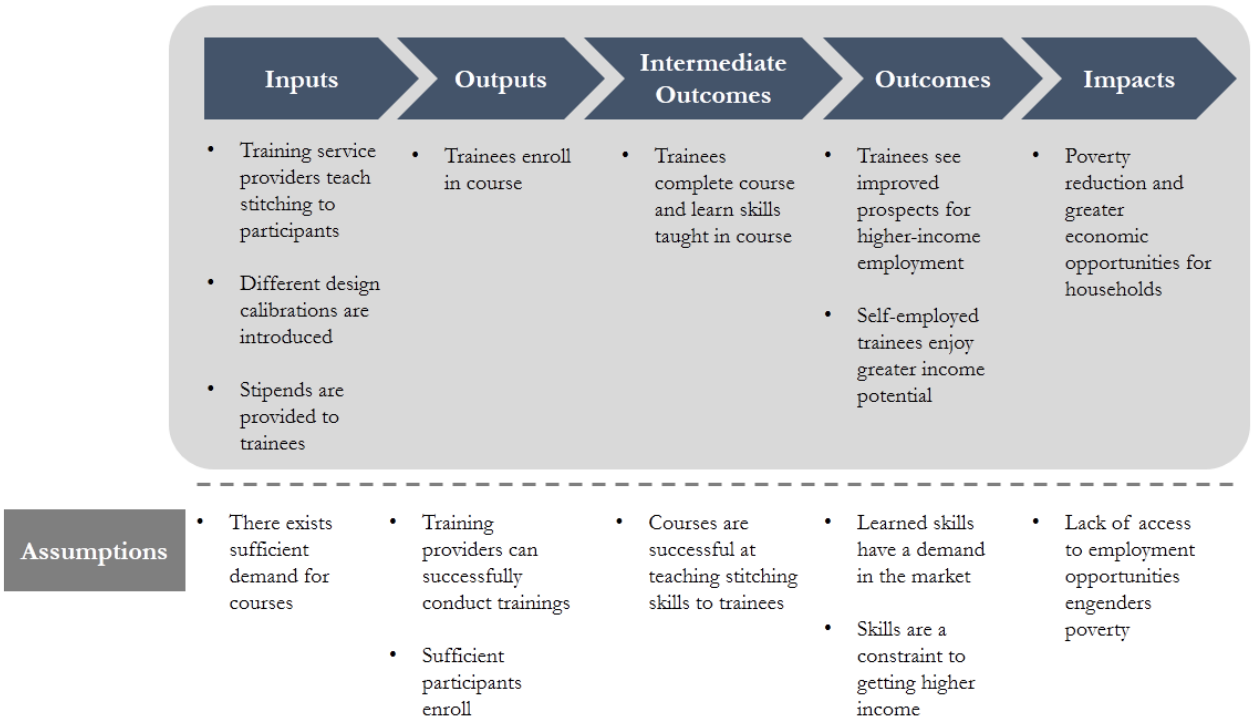
Under SFM 2013-14, we have tested the hypothesis that travel constraints are a major barrier to skills acquisition, especially among rural women and therefore alleviating them should increase uptake of skills training courses.

In order to remove the access barriers including distance, cultural norms and information deficits, PSDF introduced various design calibration in to the training scheme by either setting up training centers inside village boundary or providing group transport and conducting social mobilization activities and information sessions. The reason to do design calibrations is to figure out the primary reasons which were driving these constraints and how to best overcome them.

The following discussion outlines the theory of change:

PSDF engages the training services providers (TSPs) which conduct training and offer cash stipends to all trainees who enroll. This rests on the assumption that there exist sufficient demand for the tailoring course. TSPs, if successful in enrolling sufficient number of trainees, set up training centers and conduct trainings.. Once training courses have completed, we should observe in the immediate aftermath that trainees have learned new skills, assuming that the design calibrations were successful in eliminating access constraints and TSPs were successful at disseminating the relevant skill.
Post course completion, trainees should enjoy greater income-earning potential and improved

Figure 1: Theory of Change



prospects for gaining a higher income assuming lack of skills acts as a constraint and these skills are popularly demanded in the market. In the long run, greater access to improved employment opportunities will alleviate poverty and enhance the set of economic opportunities available to the target regions.

3 Evaluation

3.1 Scope and Objectives

PSDF's Theory of Change requires that skills training is effective at improving stitching skills and increasing income generating potential for trained women. Accordingly, our evaluation seeks to assess the impact of skills training on economic and non-economic outcomes for the rural female population of South Punjab.

The outcomes from this evaluation provide valuable insight about the impact of a female focused vocational training scheme on direct outcomes such as skills acquisition, and other downstream outcomes such as stitching, earnings, civic engagement, health, female empowerment and use of government services and employment. Moreover, we seek to measure the effects of training on citizen satisfaction using a revealed preference approach based on voting outcomes.

The evaluation studies impacts on a total of 324 villages (in which 243 were treatment villages and 81 were control villages) in the three PEOP target districts of Bahawalpur, Bahawalnagar and Muzaffargarh.

3.2 Empirical Strategy and Evaluation Design

This section outlines the evaluation strategy and design of SFM 2013-14. Section 3.2.1 discusses various design calibrations that were introduced in SFM 201-13 scheme and the motivation behind introducing them; Section 3.2.2 details the final list of treatment arms that were made part of the SFM 2013-14 scheme; Section 3.2.3 describes an Market Linkage (ML) treatment that was offered to a subset SFM 2013-14 graduates in June 2015.

3.2.1 Randomization of Intervention Variants

PSDF introduced interventions in SFM 2013-2014, designed specifically to address the access constraints that dampen rural female participation in the training schemes. These interventions were tested and implemented as different arms of the experiment. These calibrations were randomly induced allowing us to isolate their respective impacts on training outcomes.

3.2.1.1 Inducing Distance Variation

One of the best predictors of uptake in SFM 2012-13 was distance to training. To estimate the sensitivity of uptake to distance, and therefore to help with future training center allocation decisions CERP randomized villages into four groups: control; Village Based Training (VBT) in which they received a village-based training center; Group Transport (GT) in which the villagers were given support to organize reliable group transportation to and from training; and Non-VBT that were offered the add-on interventions (which are described later in the section). The random assignment of VBTs across sample villages effectively induced randomization over distance to training center within our sample. That same random assignment also effectively randomized Union Councils into receiving different numbers of training centers which enables our estimates of the training scheme's

impact on political outcomes and thus a revealed-preference assessment of citizen satisfaction with the scheme.

3.2.1.2 Inducing Variation in Safety and Reliability of Transport

CERP assisted PSDF in randomizing the villages in the Non-VBT arms into group transportation (GT), in order to address the constraints due to lack of safe and reliable transportation for women. Field teams conducted meetings with the male household members to finalize drivers and logistical arrangements of the facility, as per the suggestions given by them. These arrangements were also confirmed with the female household members to elicit their preference regarding the provision of GT. This ensured that all transport facilities including drivers, mode of transport, pick/drop locations and timings were according to the preference of the households.

3.2.1.3 Inducing Stipend Variation

In order to address financial and credit constraints and to ensure that the impact of distance and other treatment arms could be compared in a common numeraire (i.e. PKR /month), CERP offered varying stipend top-ups to trainees in the treatment villages above the standard stipend of PKR 1500 /month they received from PSDF.

To estimate the maximum range of stipend variation that could be offered **within** a village, a component of field interviews (initially conducted to inform design calibrations of SFM 2013-14scheme) was structured to elicit responses on maximum stipend difference and its acceptability. It was observed that:

- Households were more open to variation in stipend amounts if the stipend was determined through an open random ballot as they perceived that process to be fair.
- Moreover, most households expressed that a differential in stipend amounts should not exceed of PKR 1000 within a village.

In addition to the field interviews, a review of literature also supported these findings. Blount (1995), through ultimatum bargaining games in studying choice behavior, concludes that participants are more likely to accept unequal distributions when they perceive the allocation process to be fair.

Taking this into consideration, all treatment villages were randomized into 1 of 8 different stipend buckets. Within each village, stipend top-ups were then varied (into High, Medium and Low categories) at the **household level** by a random ballot. Households were explicitly informed about the procedure of allocation of stipend levels and there were no reported cases of discontentment regarding the difference in stipend values.

Table 2 list the stipend amount for each bucket and category. The stipend levels were set that for each bucket the maximum difference in stipends (i.e. difference between High and Low categories) amounted to PKR 1000 to in order to avoid a strong perception of unfairness. Table 2 also shows that the stipend in High in a lower bucket would overlap with Medium in a higher bucket, allowing us to isolate the effect due to the relative position of the stipend. Stipend assignment was stratified

across other treatment arms to ensure no significant difference in mean stipend values. This approach effectively randomized individuals into having different levels of compensation per km of distance to the nearest training center while minimizing the chances of adverse social reactions from individually randomizing stipends over a large range within each village.

Table 2: Stipend Top-up amounts

Stipend Bucket	Stipend *Top-up* Amount (PKR)		
	Low	Medium	High
1	0	500	1000
2	500	1000	1500
3	1000	1500	2000
4	1500	2000	2500
5	2000	2500	3000
6	2500	3000	3500
7	3000	3500	4000
8	3500	4000	4500

Top-up stipends for training were only offered to households in the treatment villages of evaluation sample. Within each grid, CERP randomly sorted the nine treatment villages. The first 8 were assigned to stipend buckets 1-8, and the 9th was assigned to stipend bucket 6, because previous pilots had indicated that stipends at this level were likely to maximize uptake (Cheema et al. 2013b).

3.2.1.4 Inducing Variation in Social Norms

To address the restrictive social norms, CERP helped PSDF in randomizing treatment villages in the VBT and Non-VBT arms into community mobilization. In these villages, field team conducted information sessions separately for males and females in the selected villages. Selected households from the baseline sample and respected community members were invited to attend these sessions. The purpose of these sessions was to mobilize the community and encourage household members to take part in the trainings.

3.2.2 Final Treatment Arms

The final set of treatment arms were chosen by PSDF to calibrate the effect on participation of removing different constraints as well as the cost of doing so. First was Non-VBT with standard course information only. Second was Non-VBT with standard and trainee information in which trainees were given richer information on training. Third was Non-VBT with standard information and community mobilization. Fourth was Non-VBT with standard information and group transport. Fifth was in which Non-VBT villages were given standard information, community mobilization and group transport. Sixth was basic VBT, which provided the baseline for removing the distance constraint. Seventh was VBT with standard and trainee information sessions. And the last one was VBT with community mobilization that allowed us to assess the impact of community mobilization on uptake when there is no distance constraint.

The final eight treatment arms are explained below in detail:

1. **Non-Village-based-training with Standard Information (Non-VBT):**

For Non-VBT villages, individuals did not have training centers located inside their villages but were provided with the standard information (SI) on the course content, timing and duration as well as information regarding the nearest four training centers' location. PSDF added this treatment arm to benchmark uptake rates in the other treatment arms against a context where potential trainees are provided the standard information that is provided as part of the PSDF scheme at their doorstep.

2. **Non-Village-based-training with Standard Information and Trainee Information Session (Non-VBT+TI):**

Under this treatment arm, CERP selected sample households for PSDF, which were given women-only information sessions (60 minute long) conducted by TSPs, in addition to standard course information.

3. **Non-Village-based-training with Standard Information and Community Mobilization (Non-VBT CM):**

Under this treatment arm, apart from receiving treatment 1 (Non-VBT), all-male and all-female focus groups were conducted separately with respected members of the villages in attendance, to ease social constraints. These focus groups (75 to 90 minute long) not only emphasized the potential benefits of participating in the training scheme but also examined constraints that hinder women from participating in the course. These focus groups also reflected on possible solutions to these constraints.

4. **Non-Village-based-training with Standard Information and Group transport (Non-VBT GT):**

Apart from receiving treatment 1 (Non-VBT), under this treatment, women were given an option to avail group transport to and from the training center during the course of the training, to provide safe and secure mode of transportation to women. Male members of the households were encouraged to attend meetings to suggest group transport logistics. These male members also nominated drivers responsible for the Group Transport. Preferences for Group Transport arrangements were also elicited from female trainees before finalizing them.

5. **Non-Village-based-training with Standard Information, Community Mobilization and Group Transport (Non-VBT GT & CM):** [6]

[6] Given the dual power requirements on outcomes and uptake, we could not have more than 9 treatment villages in a grid, out of which 4 had to be a VBT treatment. Our prior expectation was that community mobilization will be more effective than TI which is why the bundled intervention of Non-VBT GT and CM instead of Non-VBT GT and TI was added to establish the upper bound on training uptake among the Non-VBT arms.

In addition to receiving standard information (treatment 1) and community mobilization to ease social constraint (treatment 3), women in this treatment arm also received the option of group transport to and from the training center to ease the security constraint. Different meetings were held on the subject of Group Transport provision. Male household members were encouraged to attend and provide their input on the matters such as recruitment of drivers, routes to the training center, etc. Female trainees were also encouraged to inform their preferences regarding the logistics of Group Transport.

6. Village-based-training with Standard Information only (VBT):

Individuals in this treatment arm had training centers placed in their villages and were provided with basic course information, same as mentioned in treatment 1 (Non-VBT).

7. Village-based-training with Standard Information and Trainee Information (VBT TI):

In addition to receiving training centers in their own villages and standard information, the households selected for this treatment arm were given women-only information sessions (60 minute sessions) conducted by the TSPs that were directed to women.

8. Village-based-training with Standard Information and Community Mobilization (VBT CM):

Apart from receiving treatment 6 (VBT), two focus groups were conducted (one all-male and one all-female) with sample household members selected by CERP after randomization and with respected members of the villages. These focus groups (75 to 90 minute long) emphasized the potential benefits of participating in the training scheme and also examined constraints that hinder women from participating in the course. They also reflected collectively on possible solutions to these constraints.

Table 3 gives the summary of eight treatment arms and control group.

Table 3: Summary of Treatment Arms and Control group

	Treatment Arm	VBT	SI	TI	CM	GT
1.	Non-VBT	✗	✓	✗	✗	✗
2.	Non-VBT+TI	✗	✓	✓	✗	✗
3.	Non-VBT CM	✗	✓	✗	✓	✗
4.	Non-VBT GT	✗	✓	✗	✗	✓
5.	Non-VBT+GT + CM	✗	✓	✗	✓	✗
6.	VBT	✓	✓	✗	✗	✗
7.	VBT+ TI	✓	✓	✓	✗	✗
8.	VBT+CM	✓	✓	✗	✓	✗
9.	Control	✗	✗	✗	✗	✗

3.2.3 Market Linkages

Post treatment field visits in earlier interventions and post-treatment interviews on the SFM 2013-14 sample also highlighted a need to ease market access of trained rural women to enable them to use their newly-acquired skills. Women who had been trained reported have trouble finding places to sell their products and, hence, generate income. Because connecting SFM graduates to markets could help them utilize their skills as well as prove to be a stable source of earnings PSDF began an intervention to create market linkages. A small share of SFM 2013-14 villages in Bahawalpur and Bahawalnagar[7], therefore received this Market Linkage treatment during the period of the second follow-up tracker survey (described in Section 3.5 'Data and Surveys').[8]

3.3 Sample Size and Sample Selection

3.3.1 Evaluation Sample

There are two distinct sets of villages (called Frames) in the SFM 2013-14 training scheme. The first set—called Frame A—is the evaluation sample that will be used for understanding the impact of skills training on outcomes listed in Section 1. Frame A comprises of villages in Bahawalnagar, Bahawalpur and Muzaffargarh districts for which we have baseline data (and this set itself is a random sample of the villages in each district).

Table 4 presents the distribution of villages and households by district in Frame A.

Table 4: Number of villages and households by district

District	Number of villages	Percentage (%)	Number of Households	Percentage (%)
Bahawalnagar	120	37.04	3667	37.07
Bahawalpur	108	33.33	3302	33.38
Muzaffargarh	93	29.63	2924	29.56
Total	**324**	**100**	**9893**	**100**

These 324 villages were put into groups of twelve—called grids—based on geographical proximity. Grids were spatially spread out to ensure coverage across each of the three districts so that they represented meaningful geographical strata and were also served as useful units for randomization.

PSDF's evaluation design required a set of 3 distinct randomized assignments:

1. Allocate TSPs to **grids** such that each TSP gets assigned a random set of villages to open their centers and a set of Non-VBTs for mobilization (See Section 3.3.2)

[7] The village sample for Market Linkage treatment primarily consisted of VBTs.
[8] The impact of Market Linkages will be evaluated after the Endline Survey, as explained in Section 3.5.

2. **Within grids**, randomly assign control and treatment flavor status to each village. (See Section 3.3.3)
3. Randomly assign a stipend top-up amount to each household (See Section 3.3.2).

However, PSDF also required that each of the six TSPs has 25 VBTs for which CERP drew and additional sample of 42 villages—called Frame B—to fill this quota. **It is important to note that PSDF used Frame B villages only to fulfill its programmatic requirements; therefore, Frame B villages are excluded from the evaluation sample.**

In the following sections, the discussion will be restricted to the evaluation sample (i.e. Frame A villages)[9].

3.3.2 TSP Assignment

SFM 2013-14 required two TSPs to work within each district. PSDF assigned grids to TSPs who were selected after PSDF's rigorous selection process, to be responsible for opening and running any training centers that were assigned in those grids. In order to determine which TSP was assigned to which grid within a district, PSDF employed the following procedure:

1. CERP matched grids into pairs of two. This was done algorithmically by computing the centroid (by latitude and longitude) of each grid, and then computing the pairing that minimized the sum of distances between pairs.
2. Within pairs, grids were randomly assigned to either TSP 1 or TSP 2 of that district.

The advantage of employing the above procedure was that it ensured that TSPs had several training centers located in close proximity in order to minimize logistical challenges. The procedure also ensured that TSPs were not geographically concentrated within one part of the district.[10]

3.3.3 Village Treatment Assignment

PSDF randomly assigned villages to one of eight treatment categories or a control group within grids based on a randomization provided by CERP. Of the twelve villages per grid in Frame A, CERP assigned three to the control group, four to Village-Based-Training (VBT) treatment arms, and the remaining five to Non-Village Based-Training (Non-VBT) treatment arms. Each of the grids received a VBT with Standard Information, a VBT with Trainee Information treatment, and a VBT with Community Mobilization treatment; the remainder was either a VBT with Trainee Information treatment or a VBT with Community Mobilization treatment, which was also a result of random assignment. Table 5 below gives the total number of villages and households across treatment arms and the control group.

[9] For more information on the sampling strategy for Frame B villages, please refer to the report, 'Alleviating Access Constraints for Rural Women', (Cheema et al. 2015).

[10] This procedure also ensured that we could control for geographical and TSP variation by including grid fixed effects in the regressions.

Table 5: Number of villages and households by treatment type

Treatment Arm	Villages		Households	
	Frequency	Percentage (%)	Frequency	Percentage (%)
Non-VBT	27	8.33	851	8.6
Non-VBT+TI	27	8.33	827	8.36
Non-VBT CM	27	8.33	876	8.85
Non-VBT GT	27	8.33	895	9.05
Non-VBT+GT + CM	27	8.33	820	8.29
VBT	42	12.96	1,323	13.37
VBT+ TI	39	12.04	1,241	12.54
VBT+CM	27	8.34	887	8.97
Control	81	25	2,173	21.97
Total	**324**	**100**	**9893**	**100**

3.3.4 Power Calculations

CERP conducted power calculations to determine the sample size needed to detect uptake across treatment arms and impact on outcomes, while controlling for intra village correlation. These calculations used simulations based on the cluster-cum-household randomized design of SFM 2013-14 treatments using estimates of the average uptake and intra-cluster correlation from the earlier data available for the SFM 2012-13 evaluation. The sample size was chosen to give us at least 80% power at 5% significance level for detecting 0.2-0.3 SD impact on uptake. The sample size was designed to detect movement on socio-economic outcomes with three post-treatment rounds[11]. More survey rounds allow us to detect delayed impacts on outcomes. Two of these survey rounds are included in this report, the third will be conducted in Fall 2016. We chose to run power calculations using impact on outcomes because impact on outcomes is the limiting factor; many trainees drop out of the training at one of the four stages of uptake (voucher acceptance, voucher submission, course enrollment, course completion), leaving us a smaller sample to measure overall impacts on outcomes than to measure treatment impacts on uptake.

3.4 Treatment Balance

CERP tested whether the 8 treatment arms, and stipend treatment were balanced on household size, average ages, average schooling for women, poverty, log of per capita expenditure, assets and number of working members in the household. Balance tables in Appendix B exhibit that we are balanced on almost all key pre-treatment covariates and outcome measures. The number of unbalanced variables is expected, given our significance level of 5%. Our balance tables show that, on average, the treatment and control groups shared similar characteristics.

[11] We recognized that potentially low course uptake could leave us under powered to detect socio economic outcomes with trainees being filtered out. Having more survey rounds in these cases increases the sample size and thus increase power

3.5 Data and Surveys

We used household surveys to track our sample households. We reviewed past literature and found household surveys as the most effective way to gather data on our outcomes of interest. These surveys were done with all trainees in treatment villages as well as individuals in control villages who were identified by the head of their household as being the person who would benefit the most from vocational training. We hired a local survey firm to conduct these surveys. The survey firm hired and trained their enumerators, 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. In case the response rates were low, we drafted field strategies to address the issue of non-response.

After the completion of surveys, we were provided with the data for validity checks and cleaning before the analysis.

One baseline and two follow-up surveys are completed with our sample whereas an endline household survey is scheduled in Fall 2016.

Table 6 below list these survey rounds, their timelines and their purpose.

Table 6: Surveys conducted/planned for evaluation

Survey Round	Status	Timeline	Purpose
Baseline Survey	Completed	October to December 2013 [2-4 months prior to training]	Provided data on pre-treatment characteristics
First Follow-up Survey	Completed	December 2014 to January 2015 [4-6 months post training]	Provided data on immediate outcomes of interest post training.
Second Follow-up Survey	Completed	November 2015 to December 2015 [15-17 months post training]	Helped checking durability of these outcomes
Endline Survey	Scheduled	October to November 2016 [27-29 months post training]	Will help in fully assessing downstream impacts of skills acquisition and market linkages on household consumption, intra-household time allocation, education decisions for children, marital success, and the creation of new businesses

Additionally, CERP collected the results of 2015 local government elections from Election Commission of Pakistan. These were used for the reveled preference assessment of citizen satisfaction from the training scheme.

4 Implementation of SFM 2013-14

PSDF does not conduct trainings itself. PSDF asks training providers to submit their vocational training proposals and bid for PSDF funding. Once the training providers are selected and approved, they advertise the course, conduct mobilization activity where applicable and accept applications. Section 4.1 briefs how TSPs were hired, Section 4.2 describes their training and Section 4.3 details the rollout activities.

4.1 Hiring of Training Providers

Implementation of SFM 2013-14 required the recruitment of Training Service Providers (TSPs) that could successfully provide training in the three high poverty districts of Southern Punjab. To get shortlisted for the SFM 2013-14 scheme, all TSPs had to go through a rigorous procurement procedure as set by PSDF.

PSDF shortlisted TSPs after evaluating their Expressions of Interest (EOI) and Technical and Financial Proposals. Two TSPs per districts were shortlisted based on these proposals and the evaluation needs of the scheme. Table 7 gives the names of selected TSPs per district.

Table 7: Selected TSPs by district

District	TSP
Bahawalpur	Institute for Rural Management Kaarvan Crafts Foundation
Bahawalnagar	Aas Foundation Al-Kausar Welfare Organization
Muzaffargarh	Care Foundation Human Empowerment Foundation (HEF)

4.2 Training of Training Providers

PSDF had training sessions organized, separately, for every TSP in their respective district. Presentations were carried out to explain the demand creation strategies for different mobilization treatments. TSPs were informed about roll-out protocols and timeline details for each strategy that they had to follow as well as the evaluation needs with respect to different treatments under SFM 2013-14. PSDF also had community mobilizers trained in these sessions.

4.3 Rollout Activities

As part of the intervention, TSPs were required to disseminate information specified by village treatment type. This process was carried out in a number of visits and activities where the TSPs informed sample households about the intervention or mobilized them to take part in SFM 2013-14. Figure 2 represents different steps that needed to be carried out specific to different treatments.

Field representatives carried out **introductory visit**, common to all interventions, to inform sample households about the SFM 2013-14 training scheme. The aim of this visit was to provide sample households with information about the trainings. The representatives provided information related to course, training and stipend and also shared course booklets with the households.[12] In addition, they provided a blank enrollment form to familiarize households with the final enrollment form. Apart from this, they also invited households to attend information sessions and/or Group Transport meetings depending on their treatment arm specification.

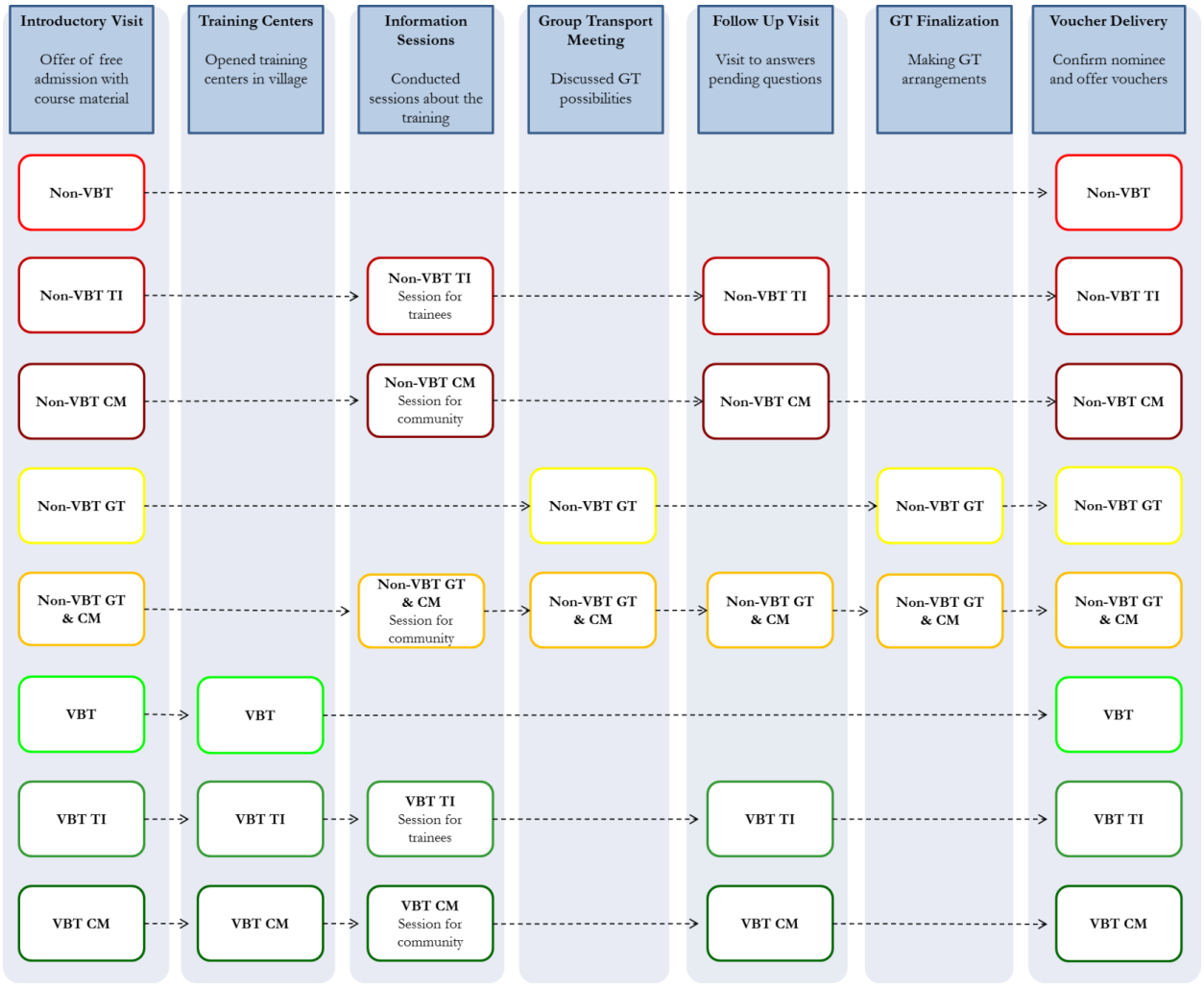
As the next step, field representatives conducted **information sessions** for treatment arms of Trainee Information (TI) and Community Mobilization (CM). For TI treatment arms, the representatives held short (60 minutes) all-female information sessions to inform eligible females regarding course content, timings and duration as well as about TSP and teacher credentials. For CM treatment arms, TSPs conducted longer (75 to 90 minutes) information sessions separately for males and females. The purpose of these sessions was to mobilize the community and encourage household members to take part in the training scheme. Like TI information sessions, course information was shared and questions relating to the scheme were answered.

After some time lapsed, the field representatives conducted **follow-up visit** for both TI and CM

Figure 2: Rollout Activities by Treatment Arms

treatment villages. They approached eligible female members from sample households to encourage them to submit vouchers. Moreover, they also addressed any pending queries that households had.

[12] The curriculum of the SFM 2013-14 was based on four months vocational skills training in domestic tailoring (adopted from TEVTA Course) with a training component on functional literacy, numeracy and financial literacy.


Introductory Visit
Offer of free admission with course material
Training Centers
Opened training centers in village
Information Sessions
Conducted sessions about the training
Group Transport Meeting
Discussed GT possibilities
Follow Up Visit
Visit to answers pending questions
GT Finalization
Making GT arrangements
Voucher Delivery
Confirm nominee and offer vouchers
Non-VBT
Non-VBT
Non-VBT TI
Non-VBT TI
Session for trainees
Non-VBT TI
Non-VBT TI
Non-VBT CM
Non-VBT CM
Session for community
Non-VBT CM
Non-VBT CM
Non-VBT GT
Non-VBT GT
Non-VBT GT
Non-VBT GT
Non-VBT GT & CM
Non-VBT GT & CM
Session for community
Non-VBT GT & CM
Non-VBT GT & CM
Non-VBT GT & CM
VBT
VBT
VBT
VBT TI
VBT TI
VBT TI
Session for trainees
VBT TI
VBT TI
VBT CM
VBT CM
VBT CM
Session for community
VBT CM
VBT CM

Rollout activities for Group Transport treatment arms required meetings to be held with male household members to discuss possible transport arrangements during the **first GT meeting**. The representatives shortlisted these arrangements after the meeting and then confirmed with female household members to elicit their preferences regarding the provision of Group Transport. They also held a **final GT meeting** with male household members in order to finalize the transportation arrangements. They provided households with printed information about the final arrangements for Group Transport facility such as driver, mode of transport, pick/drop location and timings. For the treatment arm of Non-VBT combined with GT and CM, CM information sessions and GT meetings were held separately.

Once rollout activities specific to different treatment arms had been conducted, a **voucher delivery visit** was carried out that was common to all treatment arms. During this visit, the field representatives asked sample households to nominate one eligible female member to receive training. This visit elicited voucher acceptance. If the nominee accepted the offer of course enrollment, the representative delivered vouchers to them as well.

After voucher delivery had been completed, voucher recipients were told to submit their vouchers, in case they wished to enroll, within a stipulated timeframe to the training center in which they wanted to enroll. In this **voucher submission phase**, sample households (general population) submitted vouchers at training centers where they wanted their nominated female member to get trained. Apart from sample households, self-applicants also applied - women who opted to register themselves for training in the absence of targeted information.

During the submission phase of the vouchers, it became evident that the number of applications received by TSPs exceeded class capacity. Given the number of submitted vouchers and applications along with the fact that slots were limited, CERP conducted a random ballot for PSDF to ensure that a fair and transparent allocation of slots to applicants was made without compromising the evaluation. This Enrollment Ballot determined the trainees that were shortlisted for the SFM 2013-14 scheme. As the outcome of this Enrollment Ballot, trainees were given a randomized sort order and were categorized as "Admitted" (enrolled in training) and "Waitlisted" (trainees that were kept as a backup in case admitted trainees dropped out).

TSPs announced the enrollment status of applicants for training by posting the list of admitted and waitlisted applicants at all training centers on the course start date. During this period, the field staff also visited the training center to independently record trainees' attendance. Based on these sources of information as well as TSP attendance data, applicants who did not enroll in classes post-admission lost their seats and the admission was offered to the next applicants on the waiting list (who were again informed of their new admission status by making house visits). This process continued until 95-97% of the training slots offered under SFM 2013-14 were filled after which the TSPs were allowed to fill any remaining slots on their own.

Once the class enrollment lists were finalized and communicated from TSPs to PSDF at the end of enrollment verification phase, PSDF initiated its independent monitoring process where 3rd party monitors visited each training center once a month until the course conclusion. Based on trainee attendance reports generated by these field monitoring visits (contracted out by PSDF also to a third party), PSDF determined which trainees had maintained satisfactory attendance in order to authorize their stipend payments. CERP used the same monitoring reports and attendance threshold to make the monthly stipend top-up payments to its voucher holders during the **stipend disbursement phase** of the roll out.

5 Project Timelines

The timelines of the activities for evaluation and implementation of SFM 2013-14 are illustrated in the table below.

Table 8: Project Timelines

Project Timelines	2013				2014												2015												2016											
	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Qualitative visits																																							
Baseline Survey																																							
Hiring of TSPs																																							
Training of TSPs																																							
Mobilization Activities (Community meetings, trainee info sessions, GT meetings)																																							
Course enrollment																																							
Course Training																																							
First Follow-up Survey																																							
Second Follow-up Survey																																							
Endline Survey																																							

KEY

PSDF

TSP

CERP

6 Results

This section reports key results. Section 6.1 reports on the effect of treatment type on course completion within the voucher holding population. Section 6.2 highlights the overall impact of trainings on skills-related measures and a range of downstream outcomes. Section 6.3 briefs the impact of training on household consumption and individual earnings whereas Section 6.4 assesses the impact of training on voting behavior.

6.1 Impact on Uptake

This section reports the impact of various treatment options on uptake among the voucher holder population.

Recall that to encourage prospective trainees to attend courses, TSPs offered vouchers for enrollment in the SFM 2013-14 training. Trainees who received vouchers are referred to as voucher holders; in this section, we analyze outcomes for only this voucher holding population.

We measure uptake in four stages:

- Voucher acceptance: Did the prospective trainee accept the voucher, when offered?
- Voucher submission: Did the prospective trainee submit a voucher to the preferred training center? This is analogous to submitting an application to take part in a training course.
- Course enrollment: Did the prospective trainee enroll in a course?
- Course completion: Did the prospective trainee complete the course?

Figure 3 shows the uptake rate at each stage for the three broader treatment arms: VBT, Non-VBT, and Group Transport. Uptake is calculated as a percentage of all people who were originally offered vouchers.

Figure 3: Uptake by Treatment Arm

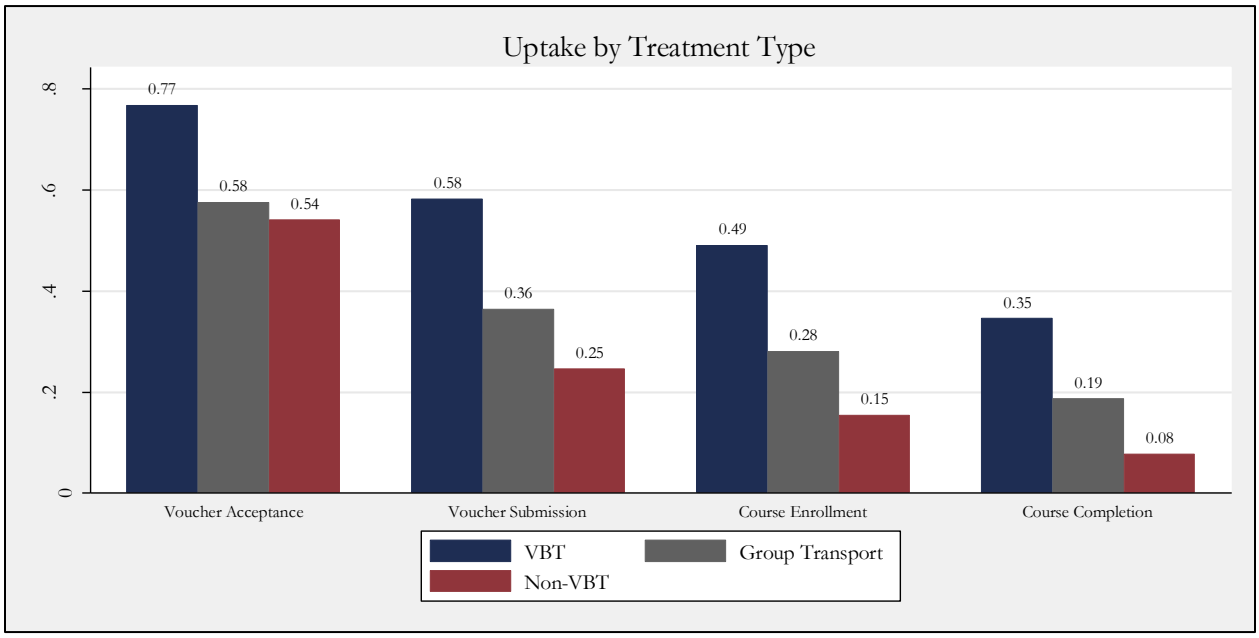


Figure 3 shows a significant drop from the voucher acceptance stage to the course completion stage for all the treatment arms. The figure also shows that uptake at each stage is highest for VBT, followed by the Group Transport and Non-VBT treatment arms.

As illustrated in the figure, 77% of prospective trainees in VBT villages accepted vouchers when offered, 75% of voucher accepters submitted their vouchers to a training center, 84% of voucher submitters enrolled in the course, and 71% of enrollees completed the course. In Group Transport villages, 58% accepted vouchers, 62% of accepters submitted vouchers, 78% of submitters enrolled in a course, and 68% of enrollees completed the course, such that only 19% of those offered vouchers completed a course. Uptake at each stage is lowest for the Non-VBT treatment arm, in which only 54% accepted vouchers and 8% of people offered vouchers completed a course.

Regression analysis for estimating the impact of various treatment arms and stipend on uptake and course completion is given in Appendix A, Table 1. The likelihood of accepting a voucher increases by 13% in the VBT arm. The probability of accepting a voucher falls by 15% for Non-VBT trainees, as compared to those who were given minimum standard information (offer of a free course and reading material) only, suggesting that providing additional information/mobilization deterred acceptance among women who knew from the information that they could not attend. For every 1000 rupees increase in the stipend, the probability of voucher acceptance increased by 3.6%.

Critically, conditional on accepting the voucher, the probability of submitting a voucher for enrollment increases by 35% for VBT treatment, as compared to the base group, and by 21% in the GT arm. The probability of voucher submission increases by 4.3% for every 1000 rupees increase in the offered stipend.

Course enrollment followed similar patterns. Conditional on submitting a voucher the probability of enrolling in SFM 2013-14 increased by 29% in VBT villages, and 22% in GT villages. Once again stipend was important, with the probability of enrollment increasing by 2.4% for every PKR 1,000/month in stipend.

VBT, GT, and stipend had similar relative impacts on course completion.

Overall, at every stage ameliorating distance constraints and easing the financial burden of acquiring training increased uptake. This is promising; the number of women targeted under SFM represents 2% of the total female population in the treatment villages and suggests scope for potential expansion of the training programs (See Appendix H for details).Full uptake results are provided in a previous CERP report, titled "Alleviating Take-up Constraints for Rural Women. (Cheema et al. 2015)"

6.2 Impact on Skill Acquisition and Downstream Outcomes

This section reports the impact of various treatment options on skill acquisition and downstream outcomes among the voucher holder population. For regression tables and details on regression specifications, refer to Appendix C. For additional outcome tables, refer to Appendix E.

For the purpose of the outcome report, we focus on the comparison between VBT and control villages. It was only in VBT villages that uptake was sufficiently high to provide a good contrast against control at this stage. Recall that total impact is defined as the product of uptake and treatment impact; insufficient uptake thus obviates analysis of impact for Non-VBT treatment arms.

For measuring the impact, IV approach has been employed to get LATE estimates, which leverages the variation in uptake induced by different treatment arms within VBT category – VBT with standard information, VBT with Trainee Information and VBT with Community Mobilization. However, it limits the results to the impact of VBT in general and does not allow providing treatment arm-specific impacts.

6.2.1 Impact on Skill Acquisition

As skills acquisition is hard to measure directly in our context we rely on a number of outcome-based measures to assess the extent of skills acquisition. If the training conferred valuable skills, then we should see that engagement in stitching activities goes up, as does the number of clothes stitched and earnings from stitching.

The variable measuring stitching activities of the individuals are defined below:

1. Engaged in Stitching (1 m) – Binary variable with value = 1 if the individual was involved in stitching during last month and value = 0 otherwise.
2. Taught Tailoring – Binary variable with value = 1 if the individual taught stitching to someone and value = 0 otherwise.

3. Stitched Clothes (during past 3 months) – Binary variable with value = 1 if the respondent stitched clothes during past 3 months and value = 0 otherwise.
4. Num. Clothes Stitched (during past 3 months) – A variable which equals to the number of clothes stitched by the individual during last 3 months.
5. Stitched for Relatives (during past 3 months) – A variable which equals to the number of clothes stitched by the individual for relatives during last 3 months.
6. Stitch Earning (Total) – Variable which equals to the earnings made by the respondent during past three months through stitching clothes
7. Stitch Earning (Non-Relative) – Variable which equals to the earnings made by the respondent during past three months through stitching clothes for non-relatives
8. Expenditure on (Tailoring Services) – Variable which equals to the expenditure on tailoring services done by the female
9. Expenditure on (Ready-Made Clothes) – Variable which equals to the amount of money spent on buying ready-made clothes

Our results (Appendix C – Table C1) show that training clearly increased stitching activity. At the first follow-up survey, we found that training increased the probability that women engaged in stitching by 10% and that they taught tailoring by 2%; by the second follow-up tracker, we find a 9% increase in probability of engagement in stitching, and a rise to 5% of increase in probability of teaching tailoring.

At the first follow-up survey, women stitched one more piece of clothing in the previous month on average. By the second follow-up survey, we saw a sharp rise in the treatment effect to 3.5 clothes in the previous month; this rise is partially explained by the rollout of the market linkages intervention, which specifically required a minimum number of stitching orders from trainees.

We see a similar result for earnings from stitching. At the first follow-up, trainees earned PKR 212 more from stitching in the previous month, of which PKR 156 came from sales to non-relatives. By the second follow-up this rose sharply to PKR 724, of which PKR 631 are attributed to sales to non-relatives. Again, the results in the second follow-up tracker are attributed to a mechanical effect of the market linkages intervention; we will find a better judgement of the effectiveness of the market linkage treatment once the endline survey is complete.

When excluding market linkages from our evaluation sample, we see a PKR 234 and PKR 221 rise in stitching earnings due to training from the first and second follow-up surveys, respectively. The effect on stitching earnings thus seems to persist at the same level for non-market linkage trainees. Extrapolating the effect on stitching earnings of non-market linkage trainees to the full 19-month period between baseline and the first follow-up, we find that trainee earnings from stitching amount to roughly PKR 4,200; critically, approximately 68% of these gains originated from sales to non-relatives.

While we had expected households to start saving on their clothing expense as a result of increased tailoring skills, we do not observe effects on tailoring expenses. This may reflect the fact that most rural women already stitched women and kids' clothes at home.

6.2.2 Impact on Downstream Outcomes

This sub-section highlights the impact of the training on downstream outcomes. We measure impacts on key downstream outcomes using the following outcome indices. (Index construction is described in detail in Appendix H and the table of regression outputs is given in Appendix C – Table C2)

- Stitching
- State Engagement
- Well Being
- Female Empowerment
- Government Services Usages
- Employment

We find that training only had an impact on the stitching index; we detected no effects on the remaining downstream outcomes. We thus see evidence that training has increased stitching activity (along with a modest increase in earnings from stitching), but has not yet impacted female empowerment, state engagement, and our remaining outcome indices. It is possible that some of our downstream outcomes will be impacted over a broader span of time – if this is the case, then we will be able to capture these effects in our planned endline survey.

We find positive correlations between course completion and both the state engagement index and the employment index. We do not find that course completion leads to increased state engagement or general employability; rather, we see that people who are more employable and more civically engaged are much more likely to complete the course. This is an important finding as it points to positive selection into training and highlights that some simple pre-screening before offering vouchers could also significantly enhance course completion.

Through a comparison of the coefficient magnitudes of Tables C1 (instrumental variables regression) and E1 (OLS regression), we see that the average woman choosing to take the course seems less oriented towards making market sales of clothes, and more towards stitching for household members and relatives. These differences highlight both the challenge and importance of designing a market access component for the kinds of rural women who can benefit most from vocational training.

6.3 Impact on Household Consumption and Individual Earnings

We also measure impact of the training on aggregate monthly household consumption and monthly individual earnings (other than stitching) using the two rounds of post-treatment survey data). We detect no change in aggregate household consumption across treated and control groups post course

completion. However, we do find positive impact on individual earnings 17 months after the end of the course.

6.4 Impact on Political Outcomes

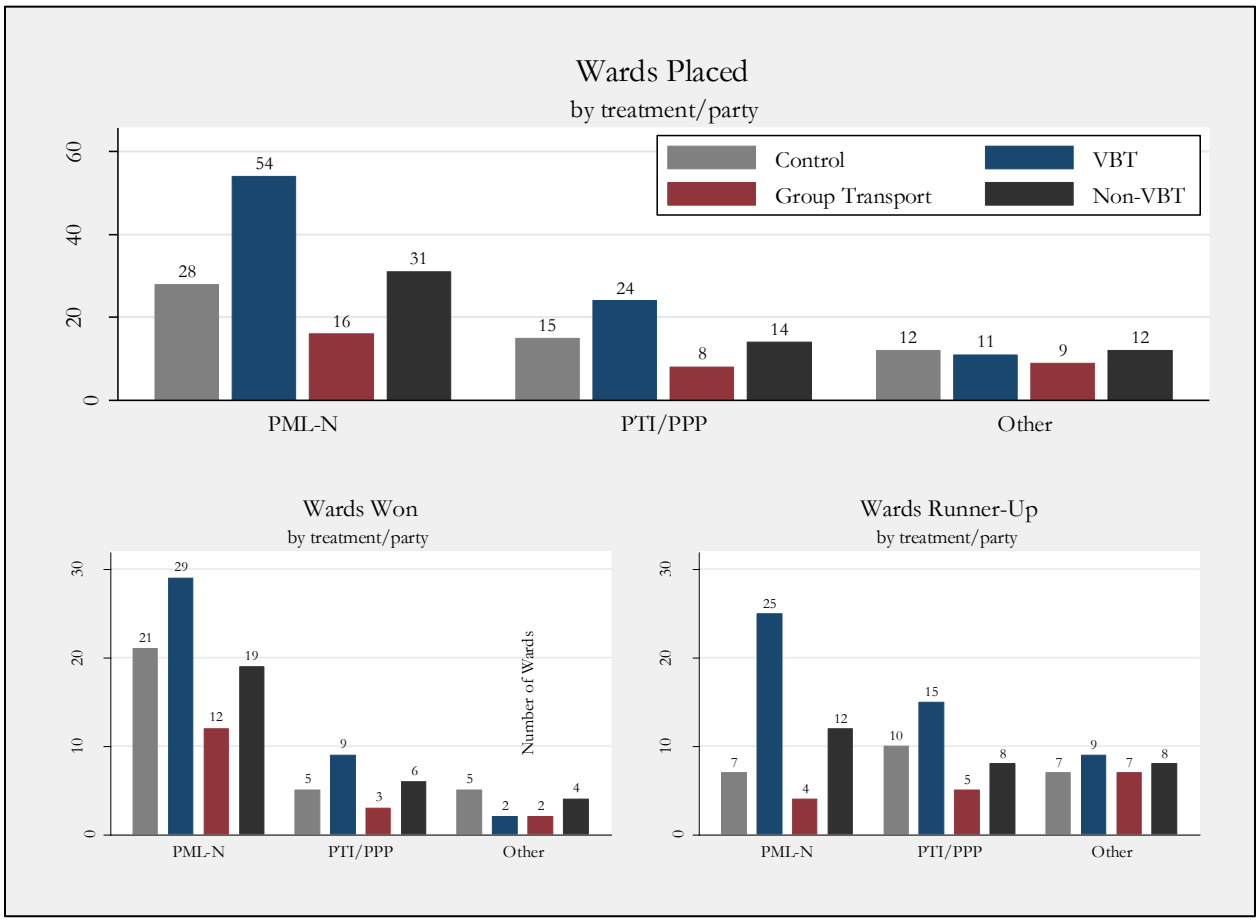
This sub-section assesses the impact of the SFM 2013-14 on aggregate voting in the Punjab Union Council Elections held in October 2015, approximately 17 months after the roll-out of SFM 2013-14. We compare election outcomes of the ruling political party Pakistan Muslim League-N (PML-N) with a group comprising of the other two well-known political parties, Pakistan Tehreek-e-Insaaf (PTI) and Pakistan Peoples' Party (PPP) using a revealed preference approach, where voting is used as a proxy for citizen satisfaction.

There are a number of issues with using subjective satisfaction with the program as an outcome variable. There is no program in control circles nor at baseline, so it is extremely artificial to ask survey respondents to rate their satisfaction with a program they have not experienced. Moreover, it is not clear that self-expressed satisfaction would be behaviorally relevant. It might be, but it's always hard to know. There is an emerging consensus in the development literature around using revealed preference measures, i.e. measures of costly behaviors where taking or refraining from an action only makes sense if preferences have shifted. In our application, the strong identification of the government with a particular party makes using impact on voting outcomes a good way to assess citizen satisfaction with the program.

As context, note that in Union Council Elections, the electoral constituency of the UC chairman and vice-chairman is the entire union council, where each union council is split in to 6 wards and comprises of multiple villages. The winner for elections (chairman and vice-chairman) is decided on basis of the aggregate votes from all six wards and multiple villages under the respective union council. Due to unavailability of voting data per candidate at the ward level, we did the analysis for the top 2 candidates at ward level (where each ward comprises of some villages which have received SFM 2013-14 treatment). Also note that we refer to a political party winning or achieving first runner up status in a ward as being "placed" in a ward.

The following Figure 4 gives the tabulation of wards placed in, by party and treatment type.

Figure 4: Number of wards placed (Winner/Runner-up)



The tabulation shows a larger percentage of villages where PML-N is placed (under all the treatment types), as compared to PTI and PPP combined.

In the control villages, PML-N is placed in 31 wards, whereas PTI/PPP is placed in 14 wards. This reflects the general popularity of the ruling political party. Among the treatment villages, PML-N is placed in 54 wards in the VBTs, as compared to PTI and PPP who are collectively placed in 24 wards in VBTs. Similarly, PML-N is placed in 16 and 31 wards as compared to PTI/PPP who are placed in 8 and 14 wards in villages given group transport and information sessions/mobilization respectively.

The division of winning outcomes shows that PML-N has won 29 wards out of 40 wards in the VBTs (≈73%), whereas PTI/PPP have collectively won 9 out of 40 wards (≈23%). Among the villages given group transport, PML-N has won 12 out of 17 wards, whereas PTI/PPP have won 3 wards. Similarly, among the villages given mobilization or information sessions (Non-VBT), PML-N has won 19 out of 29 wards and PTI/PPP have won 6 wards. These statistics show that PML-N has a larger share in the number of wards placed in (i.e. winner or runner-up) as compared to the other

two political parties among all the villages with varying assigned treatment types, with the highest one being in the VBT villages.

Our analysis shows that among the villages given VBT treatment, there is a 17% and 18% increase in the probability that a PML-N candidate will either be a winner or first runner-up of the Union Council elections (significant at 1% significance level), after controlling for the grid and political conditions. Results also show that, on top of the boost that PML-N candidates receive when their party controls the ward's constituency, they could expect an 11% increase in the probability that they would be placed in the UC elections when there was a VBT in their ward.

Similarly, in the villages given the mobilization or information session treatment (shown as Non-VBT in the above figure), a PML-N candidate can expect an increase of approximately 10% in the probability of getting placed in the ward, after netting out local political conditions. We did not detect a similar effect for Group Transport villages. (See results is Appendix D)

Overall, our analysis suggests that the ruling party's improved electoral performance was drawn not from the major rival political parties (namely, the PTI and PPP), but from independent candidates and less prominent political parties.

In a democracy, voting choices are a strong revealed preference measure of citizen satisfaction. Also as the only average difference between these villages is the treatment, hence we can infer that the strong impact on voting patterns is due to increased citizen satisfaction from village based trainings.

6.5 Methodological Limitations

Methodological limitations of using the aforementioned experimental evaluation design are listed below, along with potential mitigation strategies (where applicable):

1. Evaluation focuses on training in domestic tailoring:

This study assesses the impact of skills training in domestic tailoring on direct and downstream outcomes. Therefore, the results cannot be easily extrapolated for other types of skills trainings. However, there are a few factors that are to be considered: Domestic tailoring continues to be popular skill choice for rural women in Punjab as it enables them to work from their home which is consistent with prevailing gender norms of the region. Additionally, uptake results from PSDF's pilot interventions for rural women (Cheema et al. 2013a) showed that domestic tailoring courses had higher course completion rates than courses in home decoration and dairy products. Given female participation in vocational training schemes is generally low (as evidenced by poor course take up in the Skills for Employability scheme (Cheema et al. 2012b), PSDF was encouraged by these findings and decided to limit the course menu to just domestic tailoring in SFM 2013-14 design.

2. Evaluation is conducted on three sample districts:

This study focuses on poor and vulnerable households in the three high-poverty districts of South Punjab. While this experiment possesses internal validity, it does raise some external validity concerns. It may be the case that the results may not apply to poor and vulnerable households in other districts. However, there are a number of factors that mitigate this concern in our context. Firstly, the evaluation samples are representative of the general population and the sampling strategy ensures that our results are valid for the environment of the pilot districts. These three districts represent a large and meaningful population and the environment in these districts is similar to high poverty districts in other parts of Southern and Western Punjab and Sindh. Furthermore, the extension of the training schemes to the expansion districts allows us to draw additional representative samples of the self-selected population enrolling in training in the more developed districts and our results will be valid for the selected population in that environment. Taken together these results will provide valid insights for districts with similar populations in the province, in other provinces and in other South Asian countries.

3. Lack of voting outcomes in the baseline survey:

One issue with measuring citizen satisfaction using voting outcomes is that we only have a single snapshot of voting outcomes, captured after the intervention was completed. Since the recent Union Council elections were carried out after a pause of more than a decade, there was no historical data with which we could measure balance across treatment arms. Since we do not explicitly provide a balance table using voting outcomes, it is possible that there exist systemic differences between treatment and control groups. However, we feel that this limitation is mitigated by (a) our robust randomization procedure, and (b) sufficient balance when using all our other outcome variables.

4. Limitation of impact evaluation to VBTs:

In our evaluation, overall impact is defined as uptake multiplied by the individual treatment effect. Accordingly, low uptake hampers our attempts to detect overall impacts of training. When measuring impacts on socio-economic outcomes, we therefore limit our analysis to comparisons of VBT villages and control villages, since there is little value in measuring impacts in treatment arms where uptake is prohibitively low.

7 Conclusion

Our results show that the placement of training centers in villages (the VBT treatment) was the most successful in affecting every stage of uptake from voucher acceptance to course completion, while the next best treatment, group transportation (GT), achieves half of that effect on course completion. We also see an increase in the uptake rate as the stipend value increases, suggesting that access constraints have a monetary component.

Initial evaluation results show that training providers successfully conveyed stitching skills, as evidenced by the increase in earnings from stitching and the number of clothes stitched in the previous month. Importantly, our results show that income gains persisted between our first and second follow-up surveys (a 15-month period), and derived mostly from selling clothes to non-relatives (i.e. from small-scale market activity); this held true even for non-market linkage villages, where trainees were not given explicit minimum stitching requirements. Training thus clearly provided a monetizable skill. For nearly all downstream outcomes, however, we detect little impact.

Our results on voting outcomes – when seen as a revealed preference proxy for citizen satisfaction – suggest that citizens value the scheme. While voting outcomes are not an exact measure of citizen satisfaction, these results are encouraging for policymakers interested in increasing citizen faith in the state.

Referring to our Theory of Change, we find confirmation that some immediate outcomes (increase in stitching activity and skills) have been realized. However, impacts on longer term outcomes listed in our Theory of Change (for example, the effects on income) are not conclusive in this interim report.

Based on the interim report, we recommend making training more accessible for rural women by placing training centers in rural villages, when feasible. This will lead to an increase in the uptake and thus, a positive impact on skills acquisition, which increases stitching activity and skills level of the participants. However, given this is an interim impact evaluation report and the results are not conclusive at this stage, therefore making recommendations and discussing policy implications is not appropriate at this point. Completion of the planned endline survey will let us observe if effects persist over a longer period of time, and if certain impact effects that have not yet materialized in the first two follow-up surveys will be observed.

In addition, the endline survey will allow us to capture welfare effects of the entire household. This will also enable us to evaluate the cost-effectiveness of the intervention, which will inform PSDF's future attempts at skills interventions.

8 PEOP Log Frame

In this report, we assessed the impact of skills training in VBT villages compared to the control group as this analysis offered useful insights for the scale-up for PSDF. Additionally, we restricted the analysis to treated women and compared their outcomes against those in control group who were not offered training. This allowed us to tease-out specific value addition that training scheme may have had for the trained women (See Section 6 and Appendixes B to E for details).

However, DfID for its Program Completion Review (PCR) of PEOP is also interested in assessing the impact of mere offering the training scheme. This is called the Intent-to-Treat (ITT) effect which is simply the impact of training on all women who were offered training[13] compared to all women in the control group who were not offered training. Below we present the summary of ITT results on a range of outcome variables (See Appendix G for details):

1. Individual earnings (measured through RCT as the gap between treatment and control groups): increased by PKR 140 per month, but statistically insignificant. This represents an increase of 18.5% from the control mean.
2. Individual earnings through stitching (measured through RCT as the gap between treatment and control groups): increased by PKR 183 per month, strongly statistically significant. This represents an increase of 227% from the control mean.
3. Household consumption (measured through RCT as the gap between treatment and control groups): dropped by PKR 606 per month but statistically insignificant. This represents a 3.6% reduction from the control mean.
4. Index of wellbeing (standard deviation increase in treatment vs control individuals/households): 14 SD decrease in treatment relative to control groups and modestly statistically significant (i.e. at 10% level of significance)
5. Index of state engagement (standard deviation increase in treatment vs control individuals/households): 0.025 SD increase in treatment relative to control groups but not statistically significant.
6. Access to skills training (measured as % of women in the general population who take up PSDF training as part of CERP evaluation): 0.4% of the female population over the age of 15 years in the CERP evaluation villages, totaling 1,755 women.

[13] Women who were offered training includes all those who took up training as well those who refused the training offer.

Bibliography

Blount, S. (1995). When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences. *Organizational Behavior and Human Decision Processes*, 63(2), 131-144.

Center for Economic Research in Pakistan. (2013). SFM-B Literature Review. Center for Economic Research in Pakistan.

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

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Lodhi, A., Sheikh, S., Siddiqui, S., Tourek, G., Niazi, M., and Shoaib, A. (2012 b). The Skills for Employability Evaluation Report. Center for Economic Research in Pakistan.

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Lodhi, A., Sheikh, S., and Siddiqui, S. (2013 a). The SFM-Village Based Training Evaluation Report. Center for Economic Research in Pakistan.

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Lodhi, A., Sheikh, S., and Siddiqui, S. (2013 b). The SFJ-Stipend Evaluation Report. Center for Economic Research in Pakistan.

Cheema, A., Khwaja, A., Naseer, F., and Shapiro, J. (2014). Skills for Market 2013-2014: Design and Compliance Report. Center for Economic Research in Pakistan.

Cheema, A., Khwaja, A., Naseer, F., and Shapiro, J. (2015). Alleviating Take-up Constraints for Rural Women. Center for Economic Research in Pakistan.

Jamali, D. (2009). Constraints and opportunities facing women entrepreneurs in developing countries: A relational perspective. *Gender in Management*, 24(4), 232-251.

Kabeer, N., Huda, K., Kaur, S., and Lamhauge, N. (2012). *Productive safety nets for women in extreme poverty: Lessons from pilot projects in India and Pakistan*. Discussion Paper 28/12. London, UK: School of Oriental and African Studies, University of London.

Maitra, P., and Mani, S. (2012). *Learning and earning: Evidence from a randomized evaluation in India*

Naqvi, F. and Shahnaz, L. 2002. How Do Women Decide to Work in Pakistan?. *The Pakistan Development Review*, 41(4), Part II: 495–513.

Solotaroff, J., Hashimi, N., and Olesen, A. (2012). *Increasing Women's Employment Opportunities through TVET*. Afghanistan Gender Mainstreaming Note Series, No. 4: Technical and Vocational Education and Training (TVET). Washington, DC: The World Bank.

Wigfield, A. and Turner, R. (2012). South Asian Women and the Labour Market in the UK: Attitudes, Barriers, Solutions. *Journal of Community Positive Practices*, April 2012: 642-67.

Appendix A: Uptake Table

.

Table A1: **Effect of Treatment on Uptake**

	Voucher Acceptance	**Voucher Submission**	**Course Enrollment**	**Course Completion**
VBT	0.129***	0.346***	0.288***	0.262***
	(0.035)	(0.051)	(0.066)	(0.076)
Group Transport	-0.055	0.212***	0.221***	0.206**
	(0.043)	(0.059)	(0.071)	(0.082)
non-VBT	-0.149***	0.048	0.100	0.038
	(0.043)	(0.063)	(0.083)	(0.092)
Stipend	0.036***	0.043***	0.024***	0.047***
	(0.006)	(0.007)	(0.007)	(0.008)
Constant	0.637***	0.218**	0.456***	0.230**
	(0.064)	(0.085)	(0.096)	(0.105)
N	4034	2688	1800	1438

Notes: Each sample is conditional on completing the previous step (submission regression only includes those who accepted vouchers, etc.). The constant is the conditional mean of nVBT (Info) villages. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix B: Balance Tables

Table B1: **Balance on Pre-Treatment Outcomes**

	Control Mean	Treatment Mean	Difference of Means	p-value for Difference of Means
State Engagement:				
Social Org.	-0.002	0.002	-0.004	0.230
	(0.002)	(0.002)	(0.004)	
NGO Donation	-0.003	0.002	-0.006	0.073
	(0.002)	(0.002)	(0.003)	
Community Mediator	0.001	-0.001	0.001	0.764
	(0.004)	(0.003)	(0.005)	
Help Neighbors	-0.004	0.003	-0.006	0.481
	(0.007)	(0.006)	(0.009)	
Charity Donation	-0.024	0.017	-0.041	0.008
	(0.012)	(0.010)	(0.015)	
Total CNIC	-0.004	0.003	-0.007	0.070
	(0.003)	(0.003)	(0.004)	
Party Member	0.004	-0.003	0.007	0.009
	(0.003)	(0.001)	(0.003)	
Protest Participation	0.005	-0.003	0.008	0.006
	(0.003)	(0.002)	(0.003)	
Community Member	0.013	-0.010	0.023	0.121
	(0.011)	(0.010)	(0.015)	
Name the President	0.007	-0.005	0.013	0.149
	(0.007)	(0.006)	(0.009)	
Name the Cheif Min.	0.008	-0.006	0.014	0.229
	(0.009)	(0.008)	(0.012)	
Importance of Democracy	-0.002	0.001	-0.003	0.801
	(0.009)	(0.008)	(0.012)	
Importance of Court Independence	-0.000	0.000	-0.001	0.954
	(0.010)	(0.009)	(0.014)	
Importance of Expressing Views	0.004	-0.003	0.006	0.659
	(0.011)	(0.009)	(0.014)	
Importance of Political Involvment	0.007	-0.005	0.012	0.397
	(0.010)	(0.009)	(0.014)	
Importance of Property Rights	-0.005	0.003	-0.008	0.610
	(0.012)	(0.010)	(0.016)	
State Engagement Index	-0.060	0.044	-0.104	0.137
	(0.053)	(0.045)	(0.070)	

Table B2: **Balance on Pre-Treatment Outcomes (Cont'd)**

	Control Mean	Treatment Mean	Difference of Means	p-value for Difference of Means
Well Being:				
Physical Health	0.002	-0.002	0.004	0.717
	(0.009)	(0.008)	(0.012)	
Illness	-0.003	0.002	-0.005	0.499
	(0.006)	(0.005)	(0.007)	
Nervous	0.004	-0.003	0.007	0.361
	(0.006)	(0.005)	(0.008)	
Hopeless	0.001	-0.001	0.001	0.841
	(0.006)	(0.005)	(0.007)	
Restless	0.003	-0.002	0.005	0.525
	(0.005)	(0.005)	(0.007)	
Depresed	0.004	-0.003	0.007	0.348
	(0.005)	(0.005)	(0.007)	
Everything is an Effort	-0.003	0.002	-0.005	0.480
	(0.005)	(0.005)	(0.007)	
Worthless	0.005	-0.004	0.009	0.215
	(0.006)	(0.005)	(0.008)	
Index of Well-being	-0.053	0.038	-0.091	0.058
	(0.037)	(0.031)	(0.048)	
Employment:				
Labor Force	0.013	-0.009	0.022	0.088
	(0.010)	(0.008)	(0.013)	
Employed	0.009	-0.007	0.016	0.197
	(0.010)	(0.008)	(0.013)	
Housework	-0.010	0.007	-0.017	0.202
	(0.010)	(0.009)	(0.013)	
Self-Employ	-0.004	0.003	-0.007	0.141
	(0.003)	(0.003)	(0.005)	
Day Labor	0.005	-0.003	0.008	0.456
	(0.008)	(0.007)	(0.011)	
Employment Index	0.013	-0.010	0.023	0.348
	(0.018)	(0.016)	(0.024)	

Table B3: **Balance on Pre-Treatment Outcomes (Cont'd)**

	Control Mean	Treatment Mean	Difference of Means	p-value for Difference of Means
Stitching:				
Stitched	-0.005	0.004	-0.008	0.233
	(0.005)	(0.005)	(0.007)	
Taught Tailor	-0.001	0.001	-0.002	0.314
	(0.001)	(0.001)	(0.002)	
Stitched Clothes	-0.003	0.002	-0.004	0.404
	(0.004)	(0.004)	(0.005)	
No. Clothes	-0.008	0.006	-0.014	0.899
	(0.095)	(0.069)	(0.114)	
Stitched for Relative	-0.022	0.016	-0.038	0.397
	(0.030)	(0.031)	(0.044)	
Stitch Earnings	2.564	-1.872	4.436	0.768
	(12.898)	(8.707)	(15.010)	
Stitch Earnings (non-relative)	4.334	-3.165	7.499	0.536
	(11.758)	(5.807)	(12.129)	
Expend on Tailor	20.517	-14.980	35.497	0.311
	(27.955)	(21.882)	(35.020)	
Expend on Clothes	13.229	-9.659	22.888	0.646
	(43.130)	(28.611)	(49.793)	
Stitching Index	-0.007	0.005	-0.012	0.361
	(0.010)	(0.009)	(0.013)	
Used Services:				
Healthcare (Gov)	0.018	-0.013	0.032	0.029
	(0.011)	(0.010)	(0.014)	
Healthcare (Priv)	-0.006	0.005	-0.011	0.417
	(0.010)	(0.009)	(0.013)	
Educational	-0.002	0.002	-0.004	0.794
	(0.012)	(0.010)	(0.016)	
Police	-0.001	0.001	-0.002	0.734
	(0.005)	(0.004)	(0.006)	
Courts	-0.003	0.002	-0.005	0.357
	(0.004)	(0.003)	(0.005)	
Sanitation Service	-0.011	0.008	-0.019	0.038
	(0.007)	(0.006)	(0.009)	
Electricity	-0.001	0.001	-0.002	0.863
	(0.011)	(0.009)	(0.014)	
Services Index	-0.007	0.005	-0.011	0.782
	(0.030)	(0.027)	(0.041)	

Table B4: **Balance on Pre-Treatment Outcomes (Cont'd)**

	Control Mean	Treatment Mean	Difference of Means	p-value for Difference of Means
Female Empowerment:				
Influence Buying Land	-0.025	0.018	-0.044	0.004
	(0.012)	(0.010)	(0.015)	
Influence Borrowing	-0.029	0.021	-0.050	0.001
	(0.012)	(0.010)	(0.015)	
Permission to start Activity	-0.003	0.002	-0.006	0.576
	(0.007)	(0.007)	(0.010)	
Influence Husband's Activity	-0.026	0.019	-0.045	0.002
	(0.011)	(0.009)	(0.014)	
Influence Husband's Spending	-0.018	0.013	-0.030	0.023
	(0.010)	(0.009)	(0.013)	
Influence Daughters Education	-0.021	0.015	-0.036	0.009
	(0.011)	(0.009)	(0.014)	
Influence to Buy Sewing Machine	-0.018	0.013	-0.031	0.021
	(0.011)	(0.009)	(0.013)	
Confidence to Run Business	-0.006	0.005	-0.011	0.490
	(0.012)	(0.010)	(0.016)	
Confidence to get Credit	-0.002	0.002	-0.004	0.776
	(0.011)	(0.009)	(0.014)	
Confidence to Manage Employees	0.004	-0.003	0.007	0.605
	(0.011)	(0.009)	(0.014)	
Confidence to Manage Finances	0.000	-0.000	0.001	0.966
	(0.011)	(0.009)	(0.014)	
Confidence to Bargain	-0.003	0.002	-0.005	0.738
	(0.012)	(0.010)	(0.016)	
Confidence to Collect Debt	-0.008	0.006	-0.013	0.398
	(0.012)	(0.010)	(0.016)	
Women Manage Business	-0.018	0.013	-0.031	0.042
	(0.012)	(0.010)	(0.015)	
Girls' Education Equal	0.003	-0.003	0.006	0.700
	(0.012)	(0.010)	(0.016)	
Girls Can Engage in Paid Work	0.008	-0.006	0.014	0.192
	(0.008)	(0.007)	(0.011)	
Girls Work Outside Home	0.021	-0.015	0.036	0.012
	(0.011)	(0.009)	(0.014)	
Female Empowerment Index	-0.223	0.163	-0.386	0.001
	(0.092)	(0.078)	(0.121)	

Appendix C: Outcome Tables

Regression Specifications:

We estimate the Local Average Treatment Effect (LATE) for the training on individuals using instrumental variables (IV) regression. Because the Market Linkages (ML) treatment could impact a range of outcomes we report all results with and without the ML villages in the sample. This approach highlights the strong complementarity between creating market linkages and ket outcomes.

In each table we estimate the following:

First stages:

$$\begin{aligned} complete_i =& (VBT)\beta_1 + (VBT \times Post1)_{it}\beta_2 + (VBT \times Post2)_{it}\beta_3 \\ &+ (stipend)_i\beta_4 + (stipend \times Post1)_{it}\beta_6 + (stipend \times Post2)_{it}\beta_6 \\ &+ (Post1)_t\beta_7 + (Post2)_t\beta_8 + \epsilon \end{aligned}$$

$$\begin{aligned} (complete \times Post1)_{it} =& (VBT)_i\beta_1 + (VBT \times Post1)_{it}\beta_2 + (VBT \times Post2)_{it}\beta_3 \\ &+ (stipend)_i\beta_4 + (stipend \times Post1)_{it}\beta_6 + (stipend \times Post2)_{it}\beta_6 \\ &+ (Post1)_t\beta_7 + (Post2)_t\beta_8 + \epsilon \end{aligned}$$

$$\begin{aligned} (complete \times Post2)_{it} =& (VBT)_i\beta_1 + (VBT \times Post1)_{it}\beta_2 + (VBT \times Post2)_{it}\beta_3 \\ &+ (stipend)_i\beta_4 + (stipend \times Post1)_{it}\beta_6 + (stipend \times Post2)_{it}\beta_6 \\ &+ (Post1)_t\beta_7 + (Post2)_t\beta_8 + \epsilon \end{aligned}$$

Second stage:

$$\begin{aligned} y_{it} =& (\widehat{complete})_i\beta_1 + (\widehat{complete} \times Post1)_{it}\beta_2 + (\widehat{complete} \times Post2)_{it}\beta_3 \\ &+ (Post1)_t\beta_4 + (Post2)_t\beta_5 + \gamma + \lambda + \epsilon \end{aligned}$$

Where the “hatted" variables are the predicted values from the first stage regressions, gamma is a set of grid fixed effect dummies, lambda is stipend controls and epsilon is an error term. All results are robust to instrumenting for course enrollment as well.
In each regression the coefficient on âCompletedâ reports the mean difference between those who completed training and those who did not, ex ante at the time of baseline tracker; the coefficients on âPost 1â and âPost 2â provide the change over time in the

control group. The coefficients on âComplete x Post 1â and âComplete x Post 2â provide the causal effect of completing the training on outcomes in the first and second post-treatment surveys. At the bottom of each panel we report the average across the two post-treatment rounds and its standard error. Displaying the results this way allows a clear visualization of how the training impacts shift over time in addition to the average effect across the two rounds of follow-up survey.

Main Results:

NOTE: The sample for all regressions is limited to Voucher Holders in Frame A Villages.

Table C1: **Effect of Treatment on Stitching Outcomes**

	Engaged in Stitching (1m)	Taught Tailoring	Stitched Clothes	Num. Clothes Stitched	Stitched for Relatives	Stitch Earning (Total)	Stitch Earning (Non-Relative)	Expenditure (Tailoring Services)	Expenditure (Ready-Made Clothes)
Complete × Post 1	0.084**	0.015	0.089***	0.814*	0.394**	161.588**	103.155*	41.839	175.012
	(0.034)	(0.010)	(0.028)	(0.479)	(0.185)	(81.133)	(61.884)	(170.375)	(156.629)
Complete × Post 2	0.092***	0.050***	0.250***	3.669***	0.689***	796.613***	698.135***	-146.814	173.426
	(0.032)	(0.014)	(0.038)	(0.746)	(0.260)	(139.567)	(115.778)	(189.590)	(158.677)
Completed	0.021	0.001	0.019	0.188	0.140	6.212	-4.287	-34.856	-189.263
	(0.021)	(0.005)	(0.016)	(0.301)	(0.109)	(42.678)	(35.010)	(121.830)	(118.824)
Post 1	-0.001	0.001	0.007	0.206**	0.070*	30.871**	20.942*	-192.325***	-71.254*
	(0.006)	(0.002)	(0.005)	(0.095)	(0.037)	(13.922)	(11.487)	(41.239)	(38.605)
Post 2	-0.006	0.006***	0.009	0.244**	0.147***	46.658**	17.858	69.737	18.705
	(0.006)	(0.002)	(0.006)	(0.116)	(0.052)	(22.542)	(17.088)	(45.425)	(43.427)
Constant	0.058***	-0.005***	0.014	0.008	-0.003	-3.240	2.731	1168.729***	482.859***
	(0.014)	(0.002)	(0.012)	(0.163)	(0.059)	(31.446)	(25.371)	(58.335)	(76.906)
Post Average Effect	0.088	0.032	0.169	2.241	0.542	479.100	400.645	-52.487	174.219
	0.029	0.009	0.028	0.502	0.169	85.776	69.995	151.590	139.149
Sample Mean	0.055	0.009	0.048	0.670	0.254	109.723	79.556	768.014	462.588
Sample SD	0.229	0.093	0.214	4.502	1.897	840.812	659.404	1218.001	1418.877
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000
Panel B: Excluding ML Villages									
Complete × Post 1	0.072**	0.005	0.057*	0.701	0.273	193.452**	119.509	86.876	298.955*
	(0.033)	(0.009)	(0.031)	(0.544)	(0.211)	(95.920)	(72.878)	(193.428)	(156.583)
Complete × Post 2	0.026	0.026*	0.075**	0.855	0.168	265.664**	193.101**	-129.696	298.952*
	(0.029)	(0.015)	(0.034)	(0.567)	(0.261)	(111.033)	(76.364)	(205.646)	(166.897)
Completed	0.020	-0.001	0.033*	0.310	0.208	0.852	0.235	-96.324	-305.186***
	(0.022)	(0.006)	(0.018)	(0.331)	(0.137)	(44.032)	(33.414)	(136.635)	(117.567)
Post 1	0.000	0.002	0.009**	0.233**	0.081**	31.353**	21.525*	-198.103***	-73.550*
	(0.006)	(0.002)	(0.005)	(0.096)	(0.038)	(14.274)	(11.803)	(41.837)	(39.243)
Post 2	-0.005	0.006***	0.014***	0.374***	0.178***	71.545***	39.622***	66.896	15.680
	(0.006)	(0.002)	(0.005)	(0.107)	(0.053)	(20.873)	(14.738)	(46.289)	(44.158)
Constant	0.050***	-0.004***	0.006	-0.013	0.024	-10.353	-10.779	1166.103***	390.064***
	(0.011)	(0.001)	(0.010)	(0.162)	(0.071)	(30.558)	(18.569)	(70.462)	(60.006)
Post Average Effect	0.049	0.015	0.066	0.778	0.220	229.558	156.305	-21.410	298.953
	0.026	0.010	0.027	0.485	0.188	83.481	61.795	173.720	139.465
Sample Mean	0.050	0.007	0.040	0.545	0.239	84.191	54.853	763.120	436.741
Sample SD	0.218	0.084	0.195	4.065	1.854	774.985	585.684	1210.531	1376.293
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000

Notes: First stage regression uses VBT and stipend to instrument for completion. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table C2: **Effect of Treatment on Index Outcomes**

	Index of State Engagement	Index of Well-being	Female Empowerment Index	Gov. Services Usage Index	Employment Index	Stitching Index
Complete × Post1	0.072	0.161	0.145	0.046	-0.006	0.213***
	(0.079)	(0.108)	(0.165)	(0.102)	(0.171)	(0.075)
Complete × Post2	0.087	0.054	-0.025	0.004	-0.097	0.433***
	(0.090)	(0.095)	(0.179)	(0.098)	(0.176)	(0.076)
Completed	-0.041	-0.097	0.060	0.012	0.017	0.036
	(0.058)	(0.073)	(0.122)	(0.073)	(0.120)	(0.040)
Post 1	1.476***	0.046*	-3.140***	0.011	0.160***	0.028**
	(0.017)	(0.024)	(0.036)	(0.024)	(0.039)	(0.013)
Post 2	1.902***	0.076***	-2.914***	0.059**	0.169***	0.011
	(0.019)	(0.022)	(0.040)	(0.025)	(0.038)	(0.012)
Constant	-1.214***	-2.397***	2.141***	-1.158***	-0.420***	-0.226***
	(0.024)	(0.029)	(0.070)	(0.028)	(0.034)	(0.028)
Post Average Effect	0.079	0.107	0.060	0.025	-0.052	0.323
	0.078	0.093	0.153	0.092	0.150	0.063
Sample Mean	-0.038	-2.432	0.013	-1.026	-0.284	-0.189
Sample SD	0.907	0.466	1.689	0.445	0.838	0.492
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000
Panel B: Excluding ML Villages						
Complete × Post1	0.092	0.148	0.127	0.152	-0.076	0.162**
	(0.091)	(0.128)	(0.184)	(0.100)	(0.191)	(0.075)
Complete × Post2	0.025	0.064	-0.241	0.080	0.244	0.139**
	(0.105)	(0.105)	(0.203)	(0.098)	(0.210)	(0.070)
Completed	-0.041	-0.135	0.122	-0.026	0.045	0.043
	(0.065)	(0.085)	(0.134)	(0.074)	(0.148)	(0.043)
Post 1	1.475***	0.051**	-3.129***	0.005	0.165***	0.032**
	(0.017)	(0.025)	(0.037)	(0.025)	(0.039)	(0.013)
Post 2	1.905***	0.088***	-2.899***	0.057**	0.158***	0.019
	(0.020)	(0.022)	(0.040)	(0.025)	(0.038)	(0.012)
Constant	-1.242***	-2.421***	2.147***	-1.176***	-0.414***	-0.246***
	(0.018)	(0.033)	(0.075)	(0.031)	(0.036)	(0.024)
Post Average Effect	0.058	0.106	-0.057	0.116	0.084	0.150
	0.092	0.107	0.172	0.090	0.182	0.061
Sample Mean	-0.039	-2.440	-0.018	-1.012	-0.260	-0.207
Sample SD	0.905	0.468	1.690	0.444	0.837	0.458
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000

Notes: First stage regression uses VBT and stipend to instrument for completion. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table C3: **Effect of Treatment on HH Consumption and Individual Earnings**

	Household Consumption	Log Household Consumption	Individual Monthly Income	Log Individual Monthly Income
Complete × Post1	-1.2e+03	-0.008	-252.939	-0.089
	(1482.274)	(0.079)	(273.607)	(0.587)
Complete × Post2	-3.4e+03**	-0.117	324.537	1.268**
	(1534.879)	(0.078)	(291.450)	(0.606)
Completed	710.297	-0.012	227.101	0.838*
	(1217.389)	(0.066)	(244.458)	(0.464)
Post 1	865.281***	0.061***	-194.459***	-0.667***
	(333.232)	(0.018)	(60.084)	(0.135)
Post 2	4286.926***	0.250***	-100.949*	-0.433***
	(371.629)	(0.018)	(56.735)	(0.123)
Constant	1.7e+04***	9.626***	731.159***	1.456***
	(865.178)	(0.042)	(87.283)	(0.179)
Post Average Effect	-2.3e+03	-0.062	35.799	0.590
	1357.786	0.072	254.775	0.519
Sample Mean	1.7e+04	9.631	744.260	2.158
Sample SD	1.1e+04	0.492	1946.290	3.427
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5753.000	5753.000	5753.000	5753.000
Panel B: Excluding ML Villages				
Complete × Post1	-898.375	0.005	-412.537	-0.516
	(1533.642)	(0.083)	(324.363)	(0.630)
Complete × Post2	-3.9e+03**	-0.101	315.708	1.182*
	(1685.231)	(0.084)	(345.005)	(0.713)
Completed	919.952	-0.016	269.338	0.727
	(1303.239)	(0.071)	(297.191)	(0.531)
Post 1	690.569**	0.051***	-183.040***	-0.618***
	(327.561)	(0.017)	(61.537)	(0.137)
Post 2	4184.961***	0.241***	-111.120*	-0.427***
	(374.399)	(0.018)	(57.983)	(0.126)
Constant	1.6e+04***	9.588***	723.247***	1.377***
	(954.094)	(0.045)	(100.211)	(0.166)
Post Average Effect	-2.4e+03	-0.048	-48.414	0.333
	1463.593	0.078	309.763	0.598
Sample Mean	1.7e+04	9.625	727.867	2.054
Sample SD	1.1e+04	0.491	1975.675	3.382
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4867.000	4867.000	4867.000	4867.000

Notes: First stage regression uses VBT and stipend to instrument for completion. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix D: Voting Results

We quantify the magnitude of the effect of VBT on voting using a range of estimating approaches. The relevant rows/columns are columns (1)-(4), and the Village Based Training Row in Tables D2 and D3. Reading across the columns, we see how VBT affected PMLN outcome taking into account different factors that also effect the election. Column (1) shows the simple difference in the probability of a PMLN candidate placed across VBT and others. We estimate using the following equation:

$$wardplacement_i = (VBT)\beta_1 + (GT)\beta_2 + (non - VBT)\beta_3 + \epsilon$$

Where epsilon is an error term.
Columns (2) accounts for the randomization process by allowing for different average PMLN performance within each of the geographic grids used in assigning SFMB treatments. We estimate using the following equation:

$$wardplacement_i = (VBT)\beta_1 + (GT)\beta_2 + (non - VBT)\beta_3 + \gamma + \epsilon$$

Where gamma is a set of grid fixed effects and epsilon is an error term.
Column (3) controls for average results within the Provincial Assembly Constituencies in which the ward is located. We estimate using the following equation:

$$wardplacement_i = (VBT)\beta_1 + (GT)\beta_2 + (non - VBT)\beta_3 + \delta + \epsilon$$

Where delta is a set of Provincial Assembly Constituencies fixed effects and epsilon is an error term.
Both Columns (2) and (3) estimate the boost from placing a VBT in a ward after netting out local political conditions. Intuitively this number tells us that there can be very strong localized benefits from placing a VBT in a ward. Column (4), accounts for which party won the most recent Provincial Assembly (PA) election but assumes those effects are constant across the sample (i.e. that having a PMLN winner at the PA constituency level had the same impact on voting in the UC election for every constituency). We estimate using the following equation:

$$wardplacement_i = (VBT)\beta_1 + (GT)\beta_2 + (non - VBT)\beta_3 + \lambda + \epsilon$$

Where lambda controls for Provincial Assembly representative party and epsilon is an error term.
Intuitively, this number tells us that, on top of the boost that PMLN candidates receive when their party controls the ward's PA constituency, they could expect an 11% increase in the probability that they would place in the ward's Union Council election when there was a VBT in their ward.

Table D1: **Balance on Political Outcomes**

	Voted 2013		Voted for PMLN		Preferred Party PMLN	
Panel A: Full Treatment Breakdown						
VBT (Info)	0.063*	0.067***	0.020	0.025	0.064	0.050*
	(0.032)	(0.023)	(0.052)	(0.040)	(0.049)	(0.026)
VBT + Trainee	0.052*	0.054**	0.036	0.015	-0.018	-0.011
	(0.031)	(0.026)	(0.048)	(0.031)	(0.042)	(0.025)
VBT + Comm	0.038	0.048	0.047	0.057	0.003	-0.002
	(0.040)	(0.031)	(0.051)	(0.042)	(0.055)	(0.034)
Non-VBT (Info)	0.001	0.003	0.050	0.050	-0.015	-0.019
	(0.036)	(0.029)	(0.049)	(0.036)	(0.055)	(0.026)
Non-VBT + Trainee	0.013	0.016	-0.002	-0.005	-0.007	-0.008
	(0.034)	(0.026)	(0.073)	(0.041)	(0.057)	(0.030)
Non-VBT + Comm	-0.048	-0.045	0.043	0.048	0.004	-0.001
	(0.042)	(0.031)	(0.060)	(0.043)	(0.055)	(0.032)
Non-VBT + Transport	-0.012	-0.011	0.009	0.015	-0.020	-0.020
	(0.049)	(0.039)	(0.062)	(0.031)	(0.051)	(0.032)
Non-VBT + Transport + Comm	0.022	0.027	0.079*	0.064	0.016	0.014
	(0.035)	(0.028)	(0.047)	(0.041)	(0.055)	(0.029)
Constant	0.487***	0.419***	0.623***	0.647***	0.427***	0.644***
	(0.020)	(0.049)	(0.031)	(0.067)	(0.027)	(0.054)
Panel B: VBT vs. (Non-VBT + Control)						
VBT (Any)	0.056***	0.059***	0.010	0.008	0.022	0.019
	(0.020)	(0.015)	(0.030)	(0.022)	(0.029)	(0.017)
Constant	0.484***	0.418***	0.646***	0.668***	0.424***	0.642***
	(0.012)	(0.046)	(0.018)	(0.066)	(0.017)	(0.055)
Grid FE		X		X		X
N	8114	8114	4079	4079	8738	8738

notes: Sample limited to Frame A. Comparison Group is for Panel A is control. Grid fixed effects included where indicated. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table D2: **OLS Regression on Treatment Indicators**

	PMLN Placed				PTI/PPP Placed				Ind. Placed			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Non-VBT	0.022	0.080	0.101*	0.031	-0.006	0.007	0.010	-0.012	0.026	0.018	0.037	0.020
	(0.054)	(0.062)	(0.056)	(0.055)	(0.047)	(0.054)	(0.054)	(0.048)	(0.057)	(0.063)	(0.063)	(0.056)
Group Transport	-0.004	0.089	0.086	0.000	-0.007	-0.012	-0.015	-0.017	0.001	-0.030	0.003	0.003
	(0.062)	(0.053)	(0.051)	(0.060)	(0.044)	(0.045)	(0.056)	(0.046)	(0.062)	(0.054)	(0.068)	(0.062)
Village Based Training	0.101	0.170***	0.180***	0.106*	0.027	0.036	0.020	0.026	0.034	0.004	0.006	0.029
	(0.060)	(0.055)	(0.057)	(0.061)	(0.056)	(0.055)	(0.051)	(0.056)	(0.050)	(0.051)	(0.052)	(0.051)
Mean of Control	0.24	0.24	0.24	0.24	0.13	0.13	0.13	0.13	0.82	0.82	0.82	0.82
Regression Controls		GRID	PA	PARTY		GRID	PA	PARTY		GRID	PA	PARTY
Scaled Treat	0.010	0.496	0.521	0.048	0.002	0.145	0.149	0.026	0.002	0.239	0.258	0.031
R-squared	452	452	452	452	452	452	452	452	452	452	452	452

Notes: Sample limited to Frame A. Table reports OLS regression estimates. Placement variables are binary variables taking the value of 1 when the indicated party placed in the ward (won or was runner-up). GRID indicates Grid fixed effects. PARTY indicates that Provincial Assembly representative party is controlled for in the regression. PA indicates Provincial Assembly electorate area fixed effects. Standard errors clustered at the grid level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table D3: **OLS Regression on Treatment Indicators (Weighted)**

	PMLN Placed				PTI/PPP Placed				Ind. Placed			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Non-VBT	0.059	0.079	0.094**	0.072	-0.038	-0.030	-0.037	-0.053	0.042	0.040	0.061	0.029
	(0.049)	(0.053)	(0.044)	(0.052)	(0.043)	(0.048)	(0.043)	(0.044)	(0.048)	(0.053)	(0.055)	(0.048)
Group Transport	0.053	0.105*	0.096**	0.058	-0.049	-0.056	-0.077	-0.070	0.020	-0.000	0.034	0.021
	(0.060)	(0.055)	(0.040)	(0.059)	(0.047)	(0.048)	(0.051)	(0.046)	(0.055)	(0.050)	(0.067)	(0.059)
Village Based Training	0.146**	0.178***	0.191***	0.154***	0.015	0.022	-0.022	0.006	0.068	0.038	0.038	0.057
	(0.056)	(0.051)	(0.046)	(0.054)	(0.057)	(0.061)	(0.042)	(0.056)	(0.045)	(0.045)	(0.050)	(0.047)
Mean of Control	0.23	0.23	0.23	0.23	0.14	0.14	0.14	0.14	0.81	0.81	0.81	0.81
Regression Controls		GRID	PA	PARTY		GRID	PA	PARTY		GRID	PA	PARTY
Scaled Treat	0.014	0.465	0.536	0.057	0.004	0.095	0.171	0.030	0.005	0.237	0.272	0.040
R-squared	520	520	520	520	520	520	520	520	520	520	520	520

Notes: Sample limited to Frame A. Table reports OLS regression estimates. Placement variables are binary variables taking the value of 1 when the indicated party placed in the ward (won or was runner-up). GRID indicates Grid fixed effects. PARTY indicates that Provincial Assembly representative party controlled for in the regression. PA indicates Provincial Assembly electorate area fixed effects. Scaled treatment regressions multiply the total stipend received in the ward by the share of villages treated within the ward. Standard errors clustered at the grid level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 2 shows similar results to Table 1, but instead of treating the entire ward as having a VBT, we scale the treatment by the number of villages within the ward. For example, if a ward contains 12 villages, but only one village within that ward received a training center, then that ward receives a 1/12 treatment value instead of the 1 treatment value it would have received in Table 1. The effects estimated in Table 2 in the row labeled Village Based Training is the estimated increase in probability of placing for PMLN if every village in the ward received a training center.

Appendix E: Additional Tables

Table E1: **OLS Estimate of Effect of Treatment on Stitching Outcomes**

	Engaged in Stitching (1m)	Taught Tailoring	Stitched Clothes	Num. Clothes Stitched	Stitched for Relatives	Stitch Earning (Total)	Stitch Earning (Non-Relative)	Expenditure (Tailoring Services)	Expenditure (Ready-Made Clothes)
Complete × Post 1	0.114***	0.004	0.110***	0.984***	0.497***	122.007***	72.222**	-10.808	-35.503
	(0.017)	(0.005)	(0.015)	(0.253)	(0.109)	(40.704)	(28.020)	(52.214)	(52.930)
Complete × Post 2	0.133***	0.040***	0.246***	3.096***	0.768***	637.787***	527.784***	-60.009	-3.596
	(0.019)	(0.009)	(0.024)	(0.422)	(0.138)	(84.619)	(73.895)	(51.535)	(58.895)
Complete	0.010	0.003	0.005	0.202	0.097	21.590	13.085	-62.406	-57.309
	(0.008)	(0.003)	(0.006)	(0.158)	(0.065)	(23.542)	(15.737)	(38.915)	(42.883)
Post 1	-0.007	0.001	0.001	0.120	0.029	22.839**	17.579*	-169.204***	-44.865
	(0.005)	(0.001)	(0.004)	(0.075)	(0.030)	(11.236)	(9.015)	(33.111)	(34.084)
Post 2	-0.010**	0.008***	0.013***	0.344***	0.139***	64.661***	36.954***	74.200*	40.222
	(0.005)	(0.002)	(0.005)	(0.097)	(0.042)	(18.583)	(14.200)	(37.976)	(34.814)
Constant	0.048***	0.003***	0.024***	0.265***	0.117***	29.200***	21.364***	800.995***	484.047***
	(0.004)	(0.001)	(0.003)	(0.058)	(0.021)	(8.012)	(7.244)	(24.880)	(27.777)
Post Average Effect	0.124	0.022	0.178	2.040	0.633	379.897	300.003	-35.409	-19.549
	0.015	0.005	0.016	0.276	0.097	49.083	40.154	42.709	49.537
Sample Mean	0.055	0.009	0.048	0.670	0.254	109.723	79.556	768.014	462.588
Sample SD	0.229	0.093	0.214	4.502	1.897	840.812	659.404	1218.001	1418.877
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000
Panel B: Excluding ML Villages									
Complete × Post 1	0.098***	0.004	0.099***	0.966***	0.447***	134.460***	81.450***	25.219	-9.815
	(0.020)	(0.004)	(0.017)	(0.262)	(0.126)	(43.790)	(30.671)	(60.788)	(63.523)
Complete × Post 2	0.069***	0.029***	0.109***	1.097***	0.529***	222.901***	122.354***	-34.623	-13.928
	(0.017)	(0.009)	(0.018)	(0.286)	(0.141)	(61.723)	(43.209)	(60.675)	(66.554)
Complete	0.017*	-0.000	0.004	0.129	0.085	5.924	3.652	-70.063	-29.673
	(0.010)	(0.002)	(0.007)	(0.126)	(0.066)	(16.510)	(11.959)	(46.786)	(48.258)
Post 1	-0.005	0.002	0.000	0.138*	0.032	24.464**	18.603*	-181.196***	-37.757
	(0.005)	(0.001)	(0.004)	(0.080)	(0.033)	(12.024)	(9.721)	(35.743)	(35.333)
Post 2	-0.011**	0.006***	0.010**	0.328***	0.136***	65.131***	34.609***	62.689	49.012
	(0.005)	(0.002)	(0.005)	(0.093)	(0.045)	(18.331)	(13.109)	(39.520)	(37.269)
Constant	0.047***	0.003***	0.026***	0.266***	0.129***	27.775***	19.078***	806.068***	442.676***
	(0.004)	(0.001)	(0.003)	(0.059)	(0.024)	(7.510)	(6.446)	(26.219)	(28.126)
Post Average Effect	0.084	0.017	0.104	1.032	0.488	178.680	101.902	-4.702	-11.872
	0.016	0.005	0.014	0.232	0.106	41.930	29.195	50.544	57.238
Sample Mean	0.050	0.007	0.040	0.545	0.239	84.191	54.853	763.120	436.741
Sample SD	0.218	0.084	0.195	4.065	1.854	774.985	585.684	1210.531	1376.293
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000

Notes: Regression of outcomes on completion indicators. All regressions include grid fixed effects and stipend controls. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table E2: **OLS Estimate of Effect of Treatment on Index Outcomes**

	Index of State Engagement	Index of Well-being	Female Empowerment Index	Gov. Services Usage Index	Employment Index	Stitching Index
Complete × Post1	0.092***	0.008	0.046	0.014	0.177***	0.240***
	(0.021)	(0.027)	(0.056)	(0.027)	(0.048)	(0.035)
Complete × Post2	0.121***	-0.008	0.075	0.008	0.081	0.451***
	(0.024)	(0.024)	(0.056)	(0.027)	(0.059)	(0.048)
Completed	-0.018	0.055***	0.053	0.015	-0.061*	0.017
	(0.015)	(0.018)	(0.038)	(0.020)	(0.036)	(0.017)
Post 1	1.475***	0.053***	-3.139***	0.019	0.145***	0.017
	(0.014)	(0.019)	(0.029)	(0.019)	(0.030)	(0.011)
Post 2	1.895***	0.066***	-2.934***	0.065***	0.162***	0.017
	(0.015)	(0.017)	(0.033)	(0.019)	(0.030)	(0.011)
Constant	-1.168***	-2.473***	2.019***	-1.058***	-0.394***	-0.238***
	(0.010)	(0.013)	(0.023)	(0.014)	(0.021)	(0.008)
Post Average Effect	0.106	-0.000	0.061	0.011	0.129	0.346
	0.020	0.022	0.050	0.024	0.044	0.034
Sample Mean	-0.038	-2.432	0.013	-1.026	-0.284	-0.189
Sample SD	0.907	0.466	1.689	0.445	0.838	0.492
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5753.000	5753.000	5753.000	5753.000	5753.000	5753.000
Panel B: Excluding ML Villages						
Complete × Post1	0.089***	-0.000	0.013	0.052*	0.177***	0.219***
	(0.025)	(0.033)	(0.070)	(0.032)	(0.058)	(0.040)
Complete × Post2	0.098***	-0.000	-0.014	0.024	0.302***	0.218***
	(0.030)	(0.027)	(0.071)	(0.032)	(0.060)	(0.040)
Complete	-0.005	0.050**	0.065	0.009	-0.070	0.018
	(0.018)	(0.022)	(0.048)	(0.023)	(0.045)	(0.017)
Post 1	1.476***	0.053***	-3.127***	0.020	0.153***	0.018
	(0.015)	(0.020)	(0.031)	(0.020)	(0.033)	(0.011)
Post 2	1.891***	0.080***	-2.934***	0.070***	0.179***	0.009
	(0.017)	(0.019)	(0.035)	(0.021)	(0.032)	(0.010)
Constant	-1.165***	-2.483***	1.993***	-1.046***	-0.385***	-0.238***
	(0.010)	(0.014)	(0.024)	(0.015)	(0.023)	(0.008)
Post Average Effect	0.093	-0.000	-0.001	0.038	0.240	0.218
	0.024	0.026	0.064	0.028	0.051	0.033
Sample Mean	-0.039	-2.440	-0.018	-1.012	-0.260	-0.207
Sample SD	0.905	0.468	1.690	0.444	0.837	0.458
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4867.000	4867.000	4867.000	4867.000	4867.000	4867.000

Notes: Regression of outcomes on completion indicators. All regressions include grid fixed effects and stipend controls. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table E3: **OLS Estimate of Effect of Treatment on HH Consumption and Individual Earnings**

	Household Consumption	Log Household Consumption	Individual Monthly Income	Log Individual Monthly Income
Complete × Post1	151.693	0.010	76.332	0.538***
	(451.403)	(0.020)	(65.359)	(0.179)
Complete × Post2	-290.606	-0.029	335.243***	1.398***
	(493.808)	(0.021)	(86.493)	(0.209)
Completed	679.156*	0.054***	-48.431	0.128
	(370.304)	(0.020)	(57.584)	(0.145)
Post 1	843.875***	0.064***	-241.767***	-0.735***
	(269.533)	(0.014)	(48.445)	(0.107)
Post 2	4110.738***	0.244***	-103.630**	-0.412***
	(296.520)	(0.015)	(46.990)	(0.101)
Constant	1.6e+04***	9.522***	841.714***	2.393***
	(223.091)	(0.012)	(40.038)	(0.082)
Post Average Effect	-69.456	-0.010	205.787	0.968
	413.012	0.019	66.154	0.169
Sample Mean	1.7e+04	9.631	744.260	2.158
Sample SD	1.1e+04	0.492	1946.290	3.427
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5753.000	5753.000	5753.000	5753.000
Panel B: Excluding ML Villages				
Complete × Post1	151.397	0.019	28.600	0.407*
	(483.769)	(0.024)	(77.955)	(0.208)
Complete × Post2	-631.701	-0.030	177.917*	0.861***
	(555.059)	(0.024)	(91.386)	(0.223)
Complete	718.755*	0.044*	-69.575	-0.003
	(427.945)	(0.023)	(70.615)	(0.177)
Post 1	695.496**	0.053***	-224.862***	-0.689***
	(285.056)	(0.015)	(53.102)	(0.117)
Post 2	4044.230***	0.238***	-102.364**	-0.364***
	(320.780)	(0.015)	(50.831)	(0.111)
Constant	1.6e+04***	9.523***	829.601***	2.309***
	(238.467)	(0.012)	(43.101)	(0.088)
Post Average Effect	-240.152	-0.005	103.258	0.634
	466.348	0.022	74.323	0.191
Sample Mean	1.7e+04	9.625	727.867	2.054
Sample SD	1.1e+04	0.491	1975.675	3.382
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4867.000	4867.000	4867.000	4867.000

Notes: Regression of outcomes on completion indicators. All regressions include grid fixed effects and stipend controls. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table E4: **Intent to Treat Effect of Treatment on Stitching Outcomes**

	Engaged in Stitching (1m)	Taught Tailoring	Stitched Clothes	Num. Clothes Stitched	Stitched for Relatives	Stitch Earning (Total)	Stitch Earning (Non-Relative)	Expenditure (Tailoring Services)	Expenditure (Ready-Made Clothes)
VBT	0.011	0.000	0.009*	0.107	0.067*	3.556	-0.714	-61.039	-31.630
	(0.007)	(0.002)	(0.005)	(0.093)	(0.040)	(13.269)	(9.979)	(37.970)	(43.082)
VBT × Post1	0.013	0.000	0.012	0.028	0.019	10.152	9.397	53.256	56.421
	(0.011)	(0.003)	(0.009)	(0.138)	(0.059)	(21.493)	(15.446)	(51.516)	(54.374)
VBT × Post2	0.023**	0.016***	0.071***	0.997***	0.167**	199.063***	175.018***	-5.703	54.902
	(0.011)	(0.005)	(0.013)	(0.208)	(0.078)	(41.825)	(36.331)	(65.529)	(52.137)
Stipend	-0.002	-0.000	-0.001	-0.019	-0.009	-0.499	-0.232	25.264**	-17.389
	(0.002)	(0.001)	(0.002)	(0.030)	(0.012)	(4.093)	(3.278)	(12.310)	(14.813)
Stipend × Post1	0.008**	0.002**	0.010***	0.127**	0.059***	22.936**	13.194**	-19.180	2.506
	(0.004)	(0.001)	(0.003)	(0.054)	(0.022)	(9.409)	(6.645)	(15.358)	(18.362)
Stipend × Post2	0.004	0.001	0.008**	0.142*	0.037	39.398**	34.246**	-23.848	3.744
	(0.004)	(0.001)	(0.003)	(0.077)	(0.026)	(17.735)	(16.105)	(18.449)	(18.483)
Post 1	0.001	0.001	0.008*	0.209**	0.072**	31.784**	21.831**	-185.463***	-65.482*
	(0.006)	(0.002)	(0.004)	(0.088)	(0.034)	(12.695)	(10.522)	(38.138)	(36.037)
Post 2	-0.004	0.007***	0.016***	0.346***	0.164***	66.883***	35.674**	69.996*	24.143
	(0.005)	(0.002)	(0.005)	(0.108)	(0.049)	(20.628)	(15.392)	(41.759)	(40.195)
Constant	0.055***	-0.006***	0.008	-0.064	-0.021	-19.708	-10.845	1163.813***	482.446***
	(0.015)	(0.002)	(0.011)	(0.147)	(0.060)	(28.108)	(21.971)	(57.968)	(76.816)
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5761.000	5761.000	5761.000	5761.000	5761.000	5761.000	5761.000	5761.000	5761.000
Panel B: Excluding ML Villages									
VBT	0.011	-0.000	0.017**	0.143	0.097*	-1.394	-2.233	-62.942	-95.122**
	(0.008)	(0.002)	(0.007)	(0.101)	(0.051)	(12.778)	(9.063)	(45.408)	(40.285)
VBT × Post1	0.010	-0.001	-0.000	-0.024	-0.027	17.780	14.282	55.441	108.561**
	(0.012)	(0.003)	(0.011)	(0.153)	(0.072)	(25.026)	(17.811)	(61.964)	(54.010)
VBT × Post2	0.001	0.007	0.015	0.146	0.004	40.983	20.266	-26.762	104.819**
	(0.012)	(0.005)	(0.013)	(0.176)	(0.084)	(37.736)	(28.241)	(73.734)	(52.954)
Stipend	-0.002	0.000	-0.002	-0.013	-0.009	0.688	0.924	13.393	-8.203
	(0.003)	(0.001)	(0.002)	(0.032)	(0.014)	(4.412)	(3.386)	(12.642)	(15.733)
Stipend × Post1	0.008*	0.001	0.010***	0.129**	0.059**	24.381**	13.534*	-10.691	0.471
	(0.004)	(0.001)	(0.003)	(0.058)	(0.024)	(10.257)	(7.269)	(16.113)	(20.096)
Stipend × Post2	0.004	0.001	0.006**	0.077	0.026	25.799	23.167	-11.070	3.340
	(0.003)	(0.001)	(0.003)	(0.063)	(0.026)	(17.225)	(15.820)	(19.413)	(20.239)
Post 1	0.001	0.002	0.008*	0.219**	0.073**	31.179**	21.905**	-189.883***	-61.812*
	(0.006)	(0.002)	(0.004)	(0.089)	(0.035)	(13.067)	(10.894)	(39.032)	(37.169)
Post 2	-0.005	0.007***	0.015***	0.383***	0.176***	73.734***	39.813***	65.738	26.477
	(0.005)	(0.002)	(0.005)	(0.101)	(0.050)	(19.459)	(13.622)	(42.566)	(41.156)
Constant	0.051***	-0.004***	0.007	-0.006	0.032	-13.784	-13.529	1156.434***	379.185***
	(0.012)	(0.001)	(0.010)	(0.152)	(0.069)	(29.348)	(17.577)	(68.705)	(59.220)
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4875.000	4875.000	4875.000	4875.000	4875.000	4875.000	4875.000	4875.000	4875.000

Notes: Regression of outcomes on VBT indicators. Stipend is measured in 1000s Rps. All regressions include grid fixed. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table E5: **Intent to Treat Effect of Treatment on Index Outcomes**

	Index of State Engagement	Index of Well-being	Female Empowerment Index	Gov. Services Usage Index	Employment Index	Stitching Index
VBT	-0.008	-0.000	0.025	-0.005	-0.025	0.018
	(0.019)	(0.023)	(0.041)	(0.023)	(0.038)	(0.014)
VBT × Post1	0.021	0.018	0.014	0.022	0.008	0.031
	(0.025)	(0.034)	(0.055)	(0.034)	(0.052)	(0.023)
VBT × Post2	0.012	-0.026	-0.032	0.017	0.004	0.123***
	(0.030)	(0.030)	(0.061)	(0.031)	(0.057)	(0.027)
Stipend	-0.003	-0.017**	-0.001	0.004	0.015	-0.002
	(0.005)	(0.007)	(0.011)	(0.006)	(0.011)	(0.004)
Stipend × Post1	0.002	0.019*	0.018	-0.003	-0.005	0.022***
	(0.007)	(0.010)	(0.016)	(0.008)	(0.015)	(0.008)
Stipend × Post2	0.009	0.023**	0.010	-0.008	-0.019	0.014*
	(0.007)	(0.009)	(0.016)	(0.008)	(0.014)	(0.007)
Post 1	1.477***	0.047**	-3.137***	0.013	0.160***	0.031***
	(0.015)	(0.022)	(0.033)	(0.022)	(0.036)	(0.012)
Post 2	1.903***	0.073***	-2.916***	0.061***	0.168***	0.024**
	(0.017)	(0.020)	(0.037)	(0.023)	(0.034)	(0.012)
Constant	-1.214***	-2.396***	2.137***	-1.157***	-0.423***	-0.237***
	(0.023)	(0.028)	(0.070)	(0.027)	(0.033)	(0.027)
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5761.000	5761.000	5761.000	5761.000	5761.000	5761.000
Panel B: Excluding ML Villages						
VBT	-0.001	-0.013	0.056	-0.013	-0.024	0.023
	(0.023)	(0.028)	(0.049)	(0.025)	(0.050)	(0.016)
VBT × Post1	0.025	0.010	0.014	0.058*	0.005	0.014
	(0.031)	(0.041)	(0.067)	(0.035)	(0.064)	(0.026)
VBT × Post2	-0.017	-0.020	-0.109	0.044	0.135*	0.026
	(0.037)	(0.035)	(0.073)	(0.032)	(0.070)	(0.027)
Stipend	-0.007	-0.017**	-0.004	0.002	0.019	-0.003
	(0.005)	(0.008)	(0.012)	(0.006)	(0.012)	(0.005)
Stipend × Post1	0.004	0.020*	0.015	-0.001	-0.015	0.021**
	(0.007)	(0.012)	(0.017)	(0.009)	(0.017)	(0.008)
Stipend × Post2	0.012	0.021**	0.009	-0.007	-0.021	0.012*
	(0.008)	(0.010)	(0.017)	(0.008)	(0.016)	(0.006)
Post 1	1.477***	0.050**	-3.127***	0.012	0.166***	0.032***
	(0.016)	(0.022)	(0.034)	(0.023)	(0.036)	(0.012)
Post 2	1.902***	0.084***	-2.911***	0.062***	0.174***	0.021*
	(0.018)	(0.020)	(0.037)	(0.024)	(0.035)	(0.011)
Constant	-1.240***	-2.420***	2.151***	-1.174***	-0.420***	-0.244***
	(0.018)	(0.031)	(0.075)	(0.031)	(0.033)	(0.023)
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4875.000	4875.000	4875.000	4875.000	4875.000	4875.000

Notes: Regression of outcomes on VBT indicators. Stipend is measured in 1000s Rps. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table E6: **Intent to Treat Effect of Treatment on HH Consumption and Individual Earnings**

	Household Consumption	Log Household Consumption	Individual Monthly Income	Log Individual Monthly Income
VBT	-306.340	-0.023	90.383	0.216
	(384.466)	(0.022)	(69.290)	(0.146)
VBT × Post1	6.776	0.010	-94.974	-0.029
	(460.864)	(0.025)	(79.569)	(0.181)
VBT × Post2	-512.382	-0.021	79.759	0.402**
	(501.245)	(0.027)	(89.940)	(0.195)
Stipend	277.560**	0.010	-6.632	0.034
	(133.625)	(0.006)	(21.067)	(0.043)
Stipend × Post1	-200.724	-0.006	4.200	0.001
	(151.659)	(0.008)	(25.030)	(0.053)
Stipend × Post2	-335.893**	-0.010	18.810	0.026
	(151.544)	(0.007)	(26.334)	(0.055)
Post 1	868.350***	0.063***	-205.066***	-0.672***
	(306.332)	(0.016)	(55.169)	(0.124)
Post 2	4238.181***	0.248***	-94.825*	-0.399***
	(339.533)	(0.017)	(51.884)	(0.112)
Constant	1.7e+04***	9.627***	727.507***	1.418***
	(881.569)	(0.042)	(81.601)	(0.145)
Observations	1.7e+04	1.7e+04	1.7e+04	1.7e+04
Respondents	5761.000	5761.000	5761.000	5761.000
Panel B: Excluding ML Villages				
VBT	-278.488	-0.024	103.217	0.146
	(414.323)	(0.024)	(90.728)	(0.178)
VBT × Post1	48.946	0.011	-121.891	-0.116
	(494.488)	(0.027)	(100.975)	(0.210)
VBT × Post2	-685.130	-0.018	81.324	0.420*
	(585.489)	(0.029)	(112.924)	(0.247)
Stipend	281.408*	0.008	-2.530	0.053
	(150.335)	(0.007)	(22.772)	(0.046)
Stipend × Post1	-163.278	-0.004	-12.048	-0.030
	(165.279)	(0.008)	(27.520)	(0.056)
Stipend × Post2	-345.068**	-0.009	18.158	0.011
	(165.329)	(0.008)	(28.800)	(0.060)
Post 1	710.320**	0.052***	-196.203***	-0.630***
	(306.183)	(0.016)	(56.885)	(0.127)
Post 2	4142.118***	0.240***	-105.928**	-0.390***
	(345.723)	(0.017)	(53.203)	(0.116)
Constant	1.6e+04***	9.587***	732.898***	1.380***
	(959.544)	(0.045)	(93.770)	(0.140)
Observations	1.5e+04	1.5e+04	1.5e+04	1.5e+04
Respondents	4875.000	4875.000	4875.000	4875.000

Notes: Regression of outcomes on VBT indicators. Stipend is measured in 1000s Rps. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix F: Supplementary Tables

Table F1: **Approximate First Stage Equations**

	Completed	**Completed Post 1**	**Completed Post 2**
VBT	0.244***		
	(0.016)		
VBT × Post1		0.243***	
		(0.017)	
VBT × Post2			0.244***
			(0.017)
Stipend	0.053***		
	(0.005)		
Stipend × Post1		0.053***	
		(0.005)	
Stipend × Post2			0.053***
			(0.005)
Constant	-0.025	-0.017	-0.017
	(0.037)	(0.013)	(0.013)
N	1.7e+04	1.7e+04	1.7e+04
R-squared	0.170	0.263	0.264
F-Stat	25.231	31.030	30.767

Notes: Regression of completion status on instruments. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix G: PEOP Log Frame

Panel A - LATE IV Estimates

	Individual Earnings	Log Individual Earnings	Stitching Earnings	Log Stitching Earnings	Consumption	Log Consumption	Index of Well Being	Index of State Engagement
Course Complete × Post1	-252.94	-0.09	161.59**	0.55***	-1178.23	-0.01	0.35	0.08
	(273.61)	(0.59)	(81.13)	(0.14)	(1482.27)	(0.08)	(0.23)	(0.09)
Course Complete × Post2	324.54	1.27**	796.61***	2.03***	-3405.63**	-0.12	0.12	0.10
	(291.45)	(0.61)	(139.57)	(0.26)	(1534.88)	(0.08)	(0.20)	(0.10)
Course Complete	227.10	0.84*	6.21	0.03	710.30	-0.01	-0.21	-0.04
	(244.46)	(0.46)	(42.68)	(0.07)	(1217.39)	(0.07)	(0.16)	(0.06)
Post 1	-194.46***	-0.67***	30.87**	0.06**	865.28***	0.06***	0.10*	1.63***
	(60.08)	(0.14)	(13.92)	(0.02)	(333.23)	(0.02)	(0.05)	(0.02)
Post 2	-100.95*	-0.43***	46.66**	0.07**	4286.93***	0.25***	0.16***	2.10***
	(56.73)	(0.12)	(22.54)	(0.03)	(371.63)	(0.02)	(0.05)	(0.02)
Constant	731.16***	1.46***	-3.24	0.03	17317.87***	9.63***	0.07	-1.30***
	(87.28)	(0.18)	(31.45)	(0.07)	(865.18)	(0.04)	(0.06)	(0.03)
Sample Mean	744.26	2.16	109.72	0.27	17321.44	9.63	-0.00	0.00
Control Mean	729.95	2.02	62.46	0.14	17231.52	9.62	-0.02	-0.01
Sample SD	1946.29	3.43	840.81	1.41	10772.95	0.49	1.00	1.00
Observations	17259.00	17259.00	17259.00	17259.00	17259.00	17259.00	17259.00	17259.00
Respondents	5753.00	5753.00	5753.00	5753.00	5753.00	5753.00	5753.00	5753.00

Panel B - ITT (VBT vs. Control)

	Individual Earnings	Log Individual Earnings	Stitching Earnings	Log Stitching Earnings	Consumption	Log Consumption	Index of Well Being	Index of State Engagement
VBT × Post1	-57.71	0.04	-5.73	0.05	-118.68	0.00	0.04	0.02
	(93.06)	(0.20)	(24.77)	(0.05)	(517.28)	(0.03)	(0.08)	(0.03)
VBT × Post2	83.53	0.41*	206.61***	0.57***	-403.04	-0.01	-0.01	0.02
	(100.00)	(0.21)	(45.21)	(0.09)	(551.27)	(0.03)	(0.07)	(0.04)
VBT	110.32	0.22	9.39	0.03	-357.31	-0.03	-0.02	-0.01
	(74.52)	(0.16)	(15.67)	(0.03)	(423.30)	(0.02)	(0.05)	(0.02)
Post 1	-258.15***	-0.77***	54.41***	0.08**	1047.08**	0.07***	0.10	1.63***
	(82.04)	(0.17)	(20.02)	(0.03)	(427.34)	(0.02)	(0.06)	(0.02)
Post 2	-100.19	-0.41***	56.13*	0.11**	4082.40***	0.24***	0.09	2.10***
	(73.68)	(0.15)	(28.98)	(0.05)	(438.67)	(0.02)	(0.06)	(0.03)
Stipend × Post1	16.00	0.02	17.90*	0.06***	-240.46	-0.01	0.04*	0.00
	(27.87)	(0.06)	(10.54)	(0.02)	(160.93)	(0.01)	(0.02)	(0.01)
Stipend × Post2	20.00	0.03	41.79**	0.08***	-301.26**	-0.01	0.06***	0.01
	(28.17)	(0.06)	(18.95)	(0.03)	(151.85)	(0.01)	(0.02)	(0.01)
Stipend	0.32	0.04	1.33	-0.01	254.93*	0.01	-0.04***	-0.01
	(23.06)	(0.05)	(4.57)	(0.01)	(140.29)	(0.01)	(0.02)	(0.01)
Constant	669.84***	1.37***	-31.80	-0.01	18288.38***	9.67***	0.15**	-1.27***
	(109.14)	(0.20)	(38.12)	(0.08)	(1028.00)	(0.05)	(0.07)	(0.03)
Sample Mean	739.96	2.21	126.66	0.31	17318.35	9.63	-0.00	0.00
Control Mean	682.10	2.03	88.84	0.19	17611.89	9.65	0.00	0.00
Sample SD	1831.04	3.44	894.01	1.53	10594.17	0.49	1.01	1.00
Observations	12429.00	12429.00	12429.00	12429.00	12429.00	12429.00	12429.00	12429.00
Respondents	4143.00	4143.00	4143.00	4143.00	4143.00	4143.00	4143.00	4143.00

Notes: Regression of completion status on instruments. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

PEOP Log Frame (Cont'd)

Panel C - Treatment vs. Control	Individual Earnings	Log Individual Earnings	Stitching Earnings	Log Stitching Earnings	Consumption	Log Consumption	Index of Well Being	Index of State Engagement
Treatment × Post1	-128.43	-0.16	38.57*	0.05	265.44	0.02	0.02	0.02
	(94.90)	(0.22)	(21.99)	(0.04)	(543.07)	(0.03)	(0.08)	(0.03)
Treatment × Post2	34.70	0.20	90.47**	0.28***	-501.60	-0.02	-0.13*	0.00
	(93.65)	(0.21)	(39.92)	(0.07)	(607.68)	(0.03)	(0.08)	(0.03)
Treatment	2.27	0.10	-8.15	0.01	-53.68	0.00	0.03	0.01
	(80.53)	(0.17)	(16.46)	(0.03)	(448.59)	(0.02)	(0.06)	(0.02)
Post 1	-160.51**	-0.59***	12.79	0.05*	718.31*	0.05**	0.10	1.63***
	(74.46)	(0.18)	(15.98)	(0.03)	(434.29)	(0.02)	(0.07)	(0.03)
Post 2	-90.32	-0.39**	75.91***	0.14***	4368.95***	0.26***	0.21***	2.10***
	(72.73)	(0.16)	(29.14)	(0.04)	(504.31)	(0.02)	(0.06)	(0.03)
Stipend × Post1	16.27	0.02	17.93*	0.06***	-239.92	-0.01	0.04*	0.00
	(27.75)	(0.06)	(10.50)	(0.02)	(160.78)	(0.01)	(0.02)	(0.01)
Stipend × Post2	19.62	0.03	40.85**	0.08***	-299.42**	-0.01	0.06***	0.01
	(28.04)	(0.06)	(19.23)	(0.03)	(151.95)	(0.01)	(0.02)	(0.01)
Stipend	-0.16	0.04	0.99	-0.01	262.53*	0.01	-0.04***	-0.01
	(23.29)	(0.05)	(4.48)	(0.01)	(140.86)	(0.01)	(0.02)	(0.01)
Constant	753.21***	1.43***	-13.87	-0.01	17304.94***	9.62***	0.06	-1.30***
	(90.31)	(0.17)	(28.63)	(0.06)	(897.42)	(0.04)	(0.07)	(0.03)
Sample Mean	744.26	2.16	109.72	0.27	17321.44	9.63	-0.00	0.00
Control Mean	755.38	2.03	65.89	0.15	17329.43	9.63	0.01	-0.01
Sample SD	1946.29	3.43	840.81	1.41	10772.95	0.49	1.00	1.00
Observations	17283.00	17283.00	17283.00	17283.00	17283.00	17283.00	17283.00	17283.00
Respondents	5761.00	5761.00	5761.00	5761.00	5761.00	5761.00	5761.00	5761.00

Notes: Regression of completion status on instruments. All regressions include grid fixed effects. Standard Errors clustered at the village level reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Outcome 4 for PEOP Log Frame

SFM 2012-13	Total number of females (age>15 yrs)	48436
	Total Number of females offered vouchers	327
	% of women offered vouchers of total female population	1%
SFM 2013-14	Total number of females (age>15 yrs)	413317
	Total Number of females offered vouchers	8449
	Total number of females that enrolled	1755
	% of women offered vouchers of total female population	2%
	% of women enrolled of total female population	0.42%
	% of women enrolled of the total women who were offered vouchers	21%
SFM 2012-13 + SFM 13-14	Total number of females (age>15 yrs)	461753
	Total Number of females offered vouchers	8776
	% of women offered vouchers of total female population	2%

Appendix H:

Each index was constructed as an additive index using multiple measures, as outlined in the table below. Subtler methods of constructing indices are possible but the component-by-component results for each index are mostly insignificant. We therefore report the simplest possible summary.
Using indices instead of individual variables gives us greater power to detect effects. Consider the scenario where an overall impact on health manifests as a series of small impacts on the variables that form the Health index; in this scenario, we may be underpowered to detect small individual effects, but are better powered to detect overall effects on the Health index.
On the other hand, our use of indices obfuscates the precise channel of impact. For example, suppose we detect an impact on the Health index; in this case, it would be unclear if changes in physical health or mental health are responsible for the impact on health.

Index	Variable Name	Description
Civic Engagment	memb_socorg	Member of Social Organization
	party_member	Member of Political Party
	partic_protest	Participated in a Protest in the last 3 months
	memb_comm	Consdiers his/herself a part of the community
	name_prez	Able to name the President
	name_chiefmin	Able to name the Chief Minister
Health	nervous_m	How often they are nervous in the last month
	hopless_m	How often they feel hopless in the last month
	restless_m	How often they feel restless in the last month
	depressed_m	How often they feel depressed in the last month
	evrythgeffort_m	How often everything felt like an effort in the last month
	worthless_m	How often they felt worthless in the last month
	good_phys_health	Described physical health as good
	illness_w	Unable to perform normal activity due to illness in past 3 months
Female Empowerment	infl_delaybuyland	Ability to influence when the HH buys land
	infl_whereborrow	Ability to influence who to borrow money from
	infl_persnewactiv	Do they need to permission to start a new activity
	infl_husbnewactiv	Ability to influence husband to take up a new activity
	infl_spousespnd	Ability to influence husbad to spend more on kids clothing
	infl_girlschool	Ability to influence how long daughters stay in school
	infl_newsewmach	Ability to influence decision to buy a sewing machine
	confid_runbus	Confidence in their ability to own a business
	confid_getcredit4bus	Confidence in their ability to obtain credit for a business
	confid_mngemplys	Confidence in their ability to manage employees
	confid_mngfin	Confidence in their ability to manage finances
	confid_barg4bus_buy	Confidence in their ability to obtain cheap prices
	confid_collectdebt	Confidence in their ability to collect debts
	gender_mgmt_eq	Thinks men and women manage daily affairs equally well
	edu_eq	Thinks men and women should have equal education
	paid_work_girls	Thinks women can take paid employment
	girls_work_outside	Thinks women can work outside the home
Government Services	heagov_used	Used a governement health center in the last 3 months
	heapriv_used	Used a private health center in the last 3 months
	educ_used	Used educational services in the last 3 months
	police_used	Used police services in the last 3 months
	courts_used	Used court services in the last 3 months
	govsanit_used	Used Govt. Sanitation in the last 3 months
	elec_used	Used Electricity Company in the last 3 months
Employment	lfp_3m	Has participated in the labor force in the last 3 months
	employed_3m	Has been employed in the last 3 months
	housework_3m	Has only done housework in the last 3 months
	selfemp_3m	Has been self employed within the last 3 months
	daylab_3m	Has been a day laborer in the last three months
Stitching	stitch_enage_1m	Has engaged in stitching in the last month
	tailor_taught	Has taught tailoring
	stitch_clothes	Has stitched clothes
	no_clotes_stitched	Number of clothes stitched
	stitch_earnings	Earnings from stitching
	expend_tailor	Expenses on Tailoring
	expend_clothes	Expenses on Clothes

Appendix I

Punjab Economic Opportunities Program – Evaluation Phase Proposal

PUNJAB ECONOMIC OPPORTUNITIES PROGRAM

Evaluation Phase Proposal - CERP

Ali Cheema
Lahore University of Management Sciences

Muhammad Farooq Naseer
Lahore University of Management Sciences

Sameem I. Siddiqui
Center for Economic Research in Pakistan

Asim I. Khwaja
Harvard University

Jacob N. Shapiro
Princeton University

Ahmad Sultan
Center for Economic Research in Pakistan

TABLE OF CONTENTS

GLOSSARY

AG Agreement: Accountability Grant Agreement

Baseline Calibration: Analysis of baseline survey data that is used to inform the design effort which is being undertaken by the Punjab Skills Development Fund

BCURE: Building Capacity to Use Research Evidence.

BPSV: Big Push on Skills for Villages scheme. It aims to provide frontier skills to rural population engaged in agriculture and livestock sectors so that their productivity is enhance and poverty reduced.

Calibration: It is the estimation of some parameters of a model, under the assumption that the model is correct, as a middle step in the study of other parameters

CERP: Center of Economic Research Pakistan

CERP PIs: Center of Economic Research Pakistan Principle Investigators.

Control Group: A group of subjects in an experiment which closely resemble the treatment group in many demographic variables but not receiving the intervention/ factor under study and thereby serving as a comparison group when treatment results are evaluated.

Core Skills: Basic literacy and numeracy skills.

DFID: Department for International Development

EPoD: Evidence for Policy Design.

Evaluation: Impact evaluation with the use of RCT methodology

IDEAS: The Institute of Development and Economic Alternatives.

LEAPS: Learning and Educational Achievements of Punjab Schools.

LIC: Low Income Country

NCSW: National Commission on the Status of Women.

PBTE: Punjab Board of Technical Education.

PEOP: Punjab Economic Opportunities Program

Power Calculations: It's a calculation that informs the probability with which you can reject the hypothesis that the treatment had no effect given sample sizes, assumed baseline risk, and treatment effect size.

PSDF: Punjab Skills Development Fund

RCT: Randomized-Control-Trial; a gold standard methodology used to conduct impact evaluations (http://www.povertyactionlab.org/methodology)

RSPN: Rural Support Programs Network

Sach's Millennium Villages: A plan dedicated to ending extreme poverty in various parts of sub-Saharan Africa through targeted agricultural, medical, and educational interventions.

SFE: Skills for Employability; a training scheme introduced by PSDF in 2011-12

SFJ: Skills for Jobs scheme; a training scheme introduced by PSDF in 2012. The objective of these schemes is to increase firm productivity by addressing skills gaps faced by firms in sectors with growth and employability potential.

SFM: Skills for Markets; a training scheme introduced by PSDF in 2012. The SFM scheme has been specifically designed to provide training to the marginalized and less educated with a particular focus on women.

Smart Policy Design: A model that has evolved out of collaboration between policymakers, service providers and researchers, which cost-effectively evaluates interventions designed to address critical program challenges in their early design stages and offers the opportunity for recalibration that promises higher returns.

Treatment Group: Groups that are given the treatment and then compared with the control group to measure impact of the treatment applied.

TSP: Training Service Provider.

1. Background to the Project

Overview

This proposal outlines the activities the Center for Economic Research in Pakistan (CERP) proposes to undertake in order to deliver the impact evaluation phase of the Punjab Economic Opportunities Program (PEOP). It is in line with the collaborative agreement previously established between the Department for International Development (DFID), CERP, and the Punjab Skills Development Fund (PSDF).

In outlining the objectives, methodology, scope, associated work plan and budgetary requirements of the evaluation phase of PEOP, this proposal aims to demonstrate the value of pursing the existing collaboration between CERP, DFID and PSDF, and seeks DFID support on delivering these activities.

Background to PEOP

PEOP is the Punjab government's flagship program to catalyze inclusive growth in the Punjab province. Specifically, this proposal focuses on the evaluation of PEOP's skills program[1] managed and operated by the Punjab Skills Development Fund (PSDF).

Background on the Collaboration between CERP and DFID

The initial collaborative agreement between CERP and DFID specified that CERP's engagement would be in the following two phases:

I. Baseline Calibration and Pilot Evaluations: Provide analysis based on baseline survey data to inform the design effort which is being undertaken by the Punjab Skills Development Fund (PSDF). Concurrently pilot evaluations would be carried out of the initial PSDF schemes and learning from these evaluations will inform design. This constituted the baseline phase of the collaboration.

II. Final Program Evaluations: Conduct rigorous scientific impact evaluation using the randomized-control-trial (RCT) methodology to evaluate the impact of the final redesigned interventions designed and implemented by PSDF. This is the evaluation phase of the collaboration.

The baseline stage conducted in-depth surveys of households (31,495 households in 849 villages and urban neighborhoods) and employers (6,800 enterprises) that provided detailed insights into the demand and requirements for skills and the constraints faced in skill acquisition and employability. The evidence generated in this stage (both thorough baseline data and through three different RCT

[1] Originally PEOP consisted of skills and livestock components. However, the livestock component was discontinued in 2012.

pilot evaluations) helped PSDF to calibrate it's program design (see section 7 and Appendix D for details). Specifically, the studies conducted during the baseline phase[2] provide evidence on:

(a) What type of training best matches the demand of households and employers? This evidence[3] has been used by PSDF to program its menu of training. An important contribution of the baseline phase has been the emphasis on the importance of integrating core skills (literacy and numeracy) into the standard vocational training schemes.

(b) How should training be delivered to reduce access constraints for the target population? Rich evidence[4] on access barriers to training for different segments of the target population and a successful pilot evaluation of the *Skills for Markets* (2012-13) scheme has resulted in the introduction of village-based training as a key component of PSDF's rural skills provision menu. Similarly, the limited response shown by the urban population to stipend increases in an RCT pilot evaluation of the *Skills for Jobs* (2012-13) training scheme demonstrated the importance of job placement as a way of increasing attractiveness and returns to skills training in such environments.

(c) Which types of post-training linkages need to be strengthened to broaden the set of economic opportunities for the target population? Rich evidence[5] on high job search costs for urban citizens that arise due to narrow job search networks and high costs of access for women point to the need to integrate job placement and market linkages in scheme design.

While the initial collaborative agreement provided a budget for phase I, it was based on an assurance that both the baseline and evaluation phases will be carried out. With the successful completion of the baseline phase in the four Pilot Districts (Bahawalpur, Bahawalnagar, Lodhran and Muzzafargarh) the collaboration is now ready to transition to the evaluation phase.

---

[2] The evidence has been produced as part of six comprehensive reports that include: *The Baseline Household Report on Skills* (Cheema et. al. 2012 a); *The Baseline Household Report on Livestock* (Rasul et. al. 2012); *The Employer's Survey Report* (Cheema et. al. 2012 b); *The Skills for Employability Evaluation Report* (Cheema et. al. 2012 c); *The Skills for Market-Village Based Training Evaluation Report* (Cheema et. al. 2013 a) and *The Skills for Jobs-Stipend Report* (Cheema et. al. 2013b).

[3] This evidence has been produced as part of *The Baseline Household Survey Report on Skills* (Cheema et. al. 2012 a) and *The Employer's survey Report* (Cheema et. al. 2012b).

[4]This evidence is provided in *The Baseline Household Report on Skills* and *The Skills for Employability Evaluation Report* (*SFE)* (Cheema et. al. 2012 c), *The SFM-Village Based Training Evaluation* (Cheema et. al. 2013 a) and *The SFJ-Stipend Evaluation* (Cheema et. al 2013 b) reports

[5]This evidence is provided in *The Baseline Household Report on Skills* and *The Skills for Employability Evaluation Report* (*SFE)* (Cheema et. al. 2012 c), *The SFM-Village Based Training Evaluation* (Cheema et. al. 2013 a) and *The SFJ-Stipend Evaluation* (Cheema et. al 2013 b) reports .

2. Summary of the Evaluation Phase

Context of the PSDF Program

Punjab is undergoing a rapid demographic transition (Nayyab 2008 Amjad 2013), which has created a youth bulge that can provide an opportunity for stimulating growth but could very well turn into a disaster if this young population is unable to find productive employment. An important challenge in this context is that a large proportion of the working-age population (45% of the population in the Pilot PEOP districts) that is beyond school-going age (16-45 years old) has low human capital attainment[6]. This demographic challenge provides an important context for the design of the PEOP program.

The demographic challenge is exacerbated by the changes taking place in employment opportunities and the occupational structure in the province. The share of agriculture in total employment has declined during the past sixty years (Nasir 2003) and land-based earnings represent a shrinking opportunity set especially outside the high poverty districts. The decline in agriculture is in part due unequal land ownership, the decline in tenancy and the resumption of mechanized owner self-cultivation (Cheema and Naseer 2010). This process has led to greater reliance on non-farm employment even in the rural landless and marginal farmer population (operating less than 1 acre farms) (Malik 2005[7]) that constitute a large part of the rural poor. Low human capital attainment combined with decline in land-based opportunities is constraining a large share of the population from transitioning out of poverty and into better quality non-farm jobs and also resulting in long-run immobility of social outcomes (Cheema and Naseer 2013). **The first premise of the PEOP program's theory of change is that given this context, skills acquisition will provide critical human capital to enable the working-age population with poor human capital to transition to better quality jobs and improved earnings.** This premise is supported by preliminary evidence that acquisition of skills is positively correlated with retention of quality jobs and higher earnings (Aslam and Rawal 2013),

Another important challenge, which has increasingly gained recognition in the PSDF Board, is that stagnant productivity in the province is resulting in stagnancy of earnings in the working class population and also impeding quality job creation (Ahmed and Gautam 2013, World Bank 2012, Malik 2005). Stagnant agricultural and livestock productivity is a major factor causing poverty to persist in the high poverty districts where agriculture still dominates the occupational structure. Evidence shows that low productivity in the agriculture and livestock sectors is due to considerable dispersion of productivity across farmers operating within similar local conditions (Planning Commission 2009, Rasul et. al. 2012). This evidence suggests that productivity can be increased by widening access to best practice skills among agriculture and livestock producers within localized communities. In the case of non-agricultural sectors there is evidence of skills gaps constraining productivity with employers reporting dissatisfaction with skills in focused surveys (Cheema et. al. 2012 b, World Bank 2012). There is also evidence that demand for workers with skills and education is increasing and employers are willing to pay a premium for human capital, which suggests the

[6] Baseline evidence shows that in the pilot districts three-fifths of males and three-fourths of females in the 16-45 age-group have less than primary schooling and over two-thirds of this population falls in the poor and vulnerable (Cheema et. al. 2012 a).

[7] According to Malik's (2005) estimate rural households that are reliant for non-farm income account for over fifty percent of the rural population in the province.

existence of skills gaps (World Bank 2012). **The second premise of the program's theory of change is that widening the skills base in agriculture, manufacturing and services is essential to stimulate productivity, which is an important pre-requisite for an increase in working class earnings. Thus diffusing best practice skills within the community and providing equal access to these skills can have a significant productivity impact in the agriculture and livestock sectors, which in turn is expected to positively impact earnings in the high poverty districts. The program's theory of change also contends that access to functional training programs and resultant improvements in labor market opportunities will strengthen citizen trust in the state as well as civic and democratic political engagement. It is assumed that focusing on women as an essential target group will result in improvements in female labor force participation and gender empowerment.**

An important purpose of the evaluation phase is to rigorously evaluate PSDF's core interventions and assess if the evidence supports the program's theory of change that PSDF-supported skills acquisition will lead to better quality jobs, improved earnings, women's empowerment and improvements in citizen-trust in the state and civic and democratic political engagement.

Furthermore, evidence produced during the baseline phase showed that the expected impact of in-class training alone is likely to be constrained by certain elements of the context. The PSDF Board has used that evidence to design a number of innovative interventions that aim to improve impact by addressing these context-specific constraints. **Therefore, another important purpose of the evaluation phase is to rigorously evaluate the impact of innovative interventions designed to address context-specific constraints and produce evidence on what works and what does not and what elements of these interventions can be improved through evidence-based learning.** These innovative interventions are being designed to address the following context specific constraints:

- **Low social mobility for women:** The program is being implemented in a context of low social and labor market mobility for women. This results in three different types of constraints which are likely to reduce the impact of a program that relies solely on in-class training. Low social mobility for women increases the costs of acquiring skills training and in the absence of mitigating interventions realized demand for training is low in the wider female population. Furthermore, the demand is expected to be lower among wealth and income-constrained households that constitute the main target groups for the program. Low mobility also results in spatially narrowing the job market for women and increasing the attraction of home-based work. Therefore, training supply has to typically match opportunities available in the local job market whereby the (in-)ability of trained home-based workers to link with deeper markets becomes an important constraint to earnings. This evidence has been used by PSDF to design complementary interventions and innovations that address constraints for women related to access to training and markets. The evaluation phase will test these interventions and provide evidence for calibration.

- **Narrow job search networks:** The program is also being implemented in a context where job search networks are narrow and highly personalized. Therefore, providing the right skills may not be synonymous with getting access to quality jobs. The baseline phase has produced evidence which shows that training uptake has been low among males who are actively

engaged in the labor market because they do not expect the probability of getting quality jobs to increase after training. Evidence suggests that these pessimistic expectations result from the perception of being excluded from narrow job search networks in specific trades. This evidence has been used by the PSDF Board to design interventions to strengthen job placement and broaden access to jobs for PSDF graduates. The evaluation phase will test the impact of these interventions.

- **Unequal access to best practice skills in local communities:** The baseline evidence shows that the program is also being implemented in a context of considerable dispersion in productivity and unequal access to best practice skills among producers facing similar local conditions. This implies that positively impacting productivity and earnings may require reducing the dispersion in best practices and productivity. PSDF has used this finding to design trainings that provide access to best practice skills across communities in order to increase earnings and aggregate productivity by reducing the dispersion in farm and enterprise productivity. The evaluation phase will test the impact of this intervention as well.

These evaluations will enable us to analyze the empirical validity of a theory of change which contends that human capital acquisition is an effective channel for catalyzing growth and poverty reduction. It will do this by addressing the following questions: (1) whether acquisition of human capital is valued by employers and can help stimulate quality self-employment; (2) whether skills are demanded by households and what constraints need to be addressed to realize this demand; (3) whether skills acquisition results in better quality jobs and what types of job placement constraints need to be addressed to change perceptions that human capital acquisition is likely to result in the attainment of better quality jobs and (4) can earnings and productivity get catalyzed by providing training in a manner that diffuses skills widely within a community and exploits scale economies and complementarities at the community level.

Objectives

The objective of the final impact evaluation phase of the collaboration, given the context and its specific constraints, is t**o: (i) rigorously evaluate impact of PSDF's interventions and schemes (given below), on economic and non-economic returns in the target population of the program; (ii) provide evidence-based technical assistance for the design of innovative interventions designed to address context-specific constraints; (iii) rigorously evaluate the impact of these innovative interventions and produce evidence on what works and what does not and what elements of these interventions can be improved through evidence-based learning; (iv) create a learning space and disseminate findings through reports, peer reviewed academic publications and policy notes. ..**

PSDF schemes that will be evaluated include:

1. *Skills for Market (SFM)* scheme for rural women. The *SFM* scheme has been specifically designed to provide training to the marginalized and less educated with a particular focus on women. Therefore, evaluating the impact of this scheme for women is of great interest to PSDF as marginalized women are an important target group of the program. This evaluation will focus on rural women because a significant majority of women belonging to poor and vulnerable households reside in rural areas. In addition to examining overall impact,

calibrations designed to mitigate access constraints for rural women will also be evaluated as part of the *SFM* evaluation.

2. *Market Linkages for Skills for Market (SFM)* scheme. Given concerns that trained women in rural areas may not be able to fully leverage their skills without market access, a post-training market linkage intervention that aims to improve the returns to training for rural women will be evaluated. This will serve as a potentially critical complement to the SFM training.

3. *Skills for Jobs (SFJ)* scheme and its sectoral variants.[8] *SFJ* is the largest scheme in the PSDF portfolio and accounts for 50% of the trainings provided as part of PEOP. Therefore, the *SFJ* evaluation will be a core activity during the evaluation phase. Moreover, design calibrations designed to stimulate job matching and placement will also be evaluated as part of the *SFJ* evaluation

4. *Big Push on Skills for Villages* (*BPSV*) scheme. Over 80% of the population of the pilot districts resides in rural areas[9] and between 50-70% of the village population is engaged in agri-livestock activities (Cheema et. al. 2012). It will, therefore, be extremely difficult for PSDF to meet its poverty reduction and productivity enhancement goals without providing frontier skills to the rural population engaged in these sectors. PSDF's current portfolio is under-serving[10] the large population engaged in these sectors. The evaluation of the *BPSV* scheme being designed to address this gap will therefore be a valuable and necessary complement to PSDF's existing schemes. A novel aspect of the BPSV intervention is that it saturates villages with training in frontier skills and practices across all nodes of the value chains in agricultural, livestock and allied sectors.

Evaluation at scale will provide a powerful statistical assessment of the economic and non-economic returns associated with PSDF-supported skills training as well as the ability of the program to effectively attract and enroll members of the target population. This will be the first rigorous impact evaluation of skills training in the South-Asian context and the results will be of great relevance to PSDF, the Punjab Government, the Federal Government and multilateral and bilateral donors investing in vocational training in Pakistan and the South Asian region.

Summary of Prospective Evaluations

The baseline phase studies showed that maximizing impact requires addressing constraints along three margins: (a) ***supply*** of quality training, (b) ***access*** to training for the target population and (c) ***post-training linkages with economic opportunities*** that enable skilled workers to benefit from existing opportunities. They also show that the nature of the constraints along these margins is very different for the rural and urban populations. The rural population exists in a context of low

[8] PSDF is introducing a number of sector-specific skills schemes in the *SFJ* mode. The objective of these schemes is to increase firm productivity by addressing skills gaps faced by firms in sectors with growth and employability potential.

[9] Punjab Development Statistics 2013.

[10] Approximately 11% of people trained in PSDF-funded schemes until 30th March 2014 were trained in agriculture and livestock skills even though approximately one-third of demand for skills in the pilot districts was related to these sectors. The mismatch is due to the paucity of cost-effective providers who can supply skills for these sectors.

employer density, thin markets for training with a limited menu and severe mobility constraints. The challenges in the urban context are different as employer density is thicker as is the market for training and the menu it supports. Therefore, specific solutions to address constraints need to differ by the type of population and for each population the solutions need to focus on different margins. The four prospective evaluations will provide learning on how to solve critical constraints on impact for different populations and which margins offer the best return in each case.

Summaries of the four evaluations proposed for the evaluation phase are given below.

- **Evaluation 1 - *SFM* for rural women plus interventions to mitigate access constraints:** The *Baseline Household Report on Skills* and the *SFM-Village Based Training Evaluation* report show that **access is a critical constraint** for **rural women**. Evidence shows that in spite of strong expressed demand[11], women from the target population have a low enrollment rate[12] because of distance constraints, social barriers and constraints imposed by household responsibilities. These constraints define a social context in which women's mobility is extremely low[13] and accessing training becomes a challenge. We also find that access constraints are much more severe for women belonging to poor and vulnerable households.[14] This suggests that organizing the supply of quality training will not be enough and in order to see the desired impact, cost-effective design calibrations need to be introduced in PSDF schemes that mitigate access barriers for women and especially those belonging to poor and vulnerable households.

  The **impact evaluation of the *SFM* (2013-14) scheme (Evaluation 1) will, therefore, also evaluate the impact of interventions introduced to address access barriers.** These interventions have been designed to address specific constraints related to: distance (through the provision of village-based training and the provision of safe and reliable transport for trainees), the opportunity cost of forgoing household work and outside earnings (through stipends), social barriers (through community mobilization) and lack of information (through the provision of information on returns and opportunities).

  Evaluation 1 will provide a rigorous estimate of the economic and non-economic returns to PSDF-funded training for rural women, an essential target group of the program. It will also provide an estimate of the impact of the different design calibrations on increased access for women, in particular those belonging to poor and vulnerable households and the improvement in program returns that result from the change in access for this sub-population. This evaluation will also provide an estimate of the degree of exclusion for women belonging to poor and vulnerable households in the absence of the interventions to improve access.

- **Evaluation 2- Post-training market linkages for women trained under *SFM*:** The evidence produced during the baseline phase also shows that **post-training market linkages is a critical**

[11] The baseline survey shows that 63% of the approximately 11,000 sample households were willing or extremely willing to send a female member for training in a PSDF-supported course.
[12] The enrollment rate of women in the early PSDF offerings (*SFE*) was less than 5%.
[13] Approximately 90% of women in our baseline sample express a preference for working and training in their village and neighborhood. Moreover, less than 10% are open to working outside their village and neighborhood.
[14] The *SFE* evaluation shows that enrollment rates are 19% higher for rural non-poor as compared to the vulnerable and poor.

constraint affecting the returns to training for **rural women**. The Household Survey and Employer's Survey data shows that rural women are heavily engaged in home-based work because local employers are reluctant to hire women[15] and low labor mobility precludes migration for work. However, evidence also shows that returns to self-employment remain poor because low mobility restricts women's access to markets.

Therefore, the PSDF Board has asked CERP to **evaluate the impact of a market linkage intervention for rural women who have completed *SFM* training (Evaluation 2)**. The interest here is to introduce design calibrations to *SFM* that improve the economic returns from training for women in a context of poor access to jobs and low mobility.

Evaluation 2 will provide a rigorous estimate of the additional returns to skills training for rural women that is a result of better access to markets. This is important as the expectation is that without seeding a cost-effective post-training market linkage component the economic returns to training for women may not be substantial, which makes this a first-order margin to address.

- **Evaluation 3 – *SFJ* plus design calibrations that integrate job matching and placement:** Evidence from the baseline phase shows that **post-training job placement** is a critical constraint for the urban population. Evidence from the baseline phase shows that for the urban population PSDF has been successful in seeding the supply of accredited trainings in a broad menu of trades and as a result access is not a big issue. However, the evidence also shows that the expectation of getting skilled jobs after training remains low among urban trainees, which results in low enrollment rates among individuals who are actively participating in the labor market[16] (Cheema et. al. 2013 b). Qualitative interviews reveal that potential trainees are unwilling to forgo their present jobs to enroll in a course because they remain pessimistic about getting a suitable job after training. This is a consequence of highly personalized and narrow job search networks that raise information costs for both employers and employees and increases search costs on both sides of the market.

  This evidence has created recognition in the PSDF Board that it will be difficult to ensure adequate returns to training for the urban population without complementing in-class training with a post-training job matching and placement service component**.** Therefore, **in addition to the standard *SFJ* evaluation**, the PSDF Board has requested **an evaluation of a pilot with a complementary job matching and placement component (Evaluation 3)**.

  Evaluation 3 will provide a rigorous estimate of the economic and non-economic returns to training for the urban population benefitting from PSDF-funded high end training. It will also allow us to detect the additional impact of complementary job matching and placement components on those who attend training and the impact on enrollment in the general population, in particular among active labor market participants.

- **Evaluation 4 -*BPSV* for the rural population:** The baseline phase evidence shows that **access to the right menu of training and a mode of delivery that exploits complementarities are**

[15] The Employer's survey shows that over 90% of enterprises are exclusively male and do not employ women.
[16] The enrollment rate among active male labor market participants is less than 10%, which is much lower than the enrollment rate for students (22%) and the unemployed (13%). This is worrying as the baseline data shows that over 80% of males are actively engaged in the labor market.

critical constraints for the broader rural population. The baseline reports identify supply-demand mismatches for the rural population: Approximately one-third of the households in the pilot districts would like to acquire skills related to the agri-livestock sectors and but only around 11% of PSDF-supported graduates have acquired related skills. Therefore, increasing the supply of training for these sectors remains an important challenge for the program. The PSDF Board, however, recognizes that doing so effectively requires an innovative model of agri-livestock training for villages that both introduces a comprehensive menu of context-relevant training in frontier practices and skills through a credible consortium that has demonstrated success and provides these skills to a large fraction of community members and along the entire value chains. The model is motivated by evidence that whereas progressive farmers and private agro-livestock companies have shown great success utilizing frontier skills and practices in the pilot districts, the majority of farmers in adjacent or similar areas have proven incapable of adoption. What is startling is that progressive farmers that are utilizing frontier practices are achieving up to three times higher yield[17] (Ahmed and Gautam 2013, Rasul et. al. 2013). The model is also motivated by literature that shows low returns to incremental investments in skills and technologies because they fail to take advantage of scale economies and production complementarities (Murphy et. al. 1989, Kremer 1993). There is recognition that large-scale impact may be difficult without exploiting scale economies and production complemenarities, which can be achieved through a model that saturates the community with training across all relevant nodes of the value chain.

The PSDF Board has asked CERP to **evaluate the impact of BPSV for the broader rural population (Evaluation 4)**. There is great interest in learning about the potential of introducing an innovative model of training that is relevant for the broader rural population, which draws on frontier skills and practices and takes advantage of scale economies and production complementarities in communities to realize the greatest benefit for a given cost.[18]

Evaluation 4 will rigorously measure the impact of intensive human capital infusion through training in frontier agri-livestock skills on economic and non-economic returns at the household and village levels. The existence of a large household baseline sample, which allows the estimation of village-level measures of income, wealth and capital stock, has created the novel opportunity to measure impact on village-level GDP and productivity, gains to households at different points in the wealth distribution and the extent of spillovers that result from this intervention. This evaluation will be able to provide the impact of training on aggregate productivity and poverty reduction at the community level.

The evaluations, along with the calibration approach, were approved by the PSDF Board in the 17th meeting of its Board, which was held on the 7th of February 2014 (Appendix B gives the text of the relevant item viii from the minutes). The four prospective evaluations will maximize the use of the extensive baseline data that has been collected for a representative sample of households in the pilot districts during the baseline phase.

---

[17] The *CERP Baseline Household Report on Livestock* finds an enormous variation in milk yields in the same village with the 25% most productive households having productivity levels that are double those of the least 25% productive households.

[18] Similar ideas underlie the Millennium Villages model championed by Jeffery Sachs in over 20 countries (Sachs 1999).

The skills component of the PEOP program has been (and will continue to be) rolled out in the four high-poverty "pilot districts" in South Punjab. However, during 2012-13 PEOP was redesigned as a stand-alone skills program that is in the process of being scaled-up in 10 additional districts, which are referred to as the "expansion districts".[19] The main focus of the evaluation phase will be on the pilot PEOP districts, which represent the flagship of the current program. Extensive baseline data has already been collected and analyzed for the pilot districts, and as a result, key program challenges have already been identified. Although the Evaluation Phase focuses on the pilot districts, learning will also take place within the expansion districts. This is because a number of the expansion districts have economies that are very similar to the pilot districts[20]. In addition, the evaluation of PSDF's main *SFJ* scheme and its sectoral variants will be conducted in a sample of the pilot and expansion districts.

The evaluation phase of this agreement will continue to follow the collaborative arrangement between the CERP PIs and DFID that was followed in the baseline phase. The four CERP PIs will contribute the cost of their time as part of this collaborative arrangement and DFID will support the costs related to interventions, surveys[21] and dissemination.

Work Plan

The evaluation phase proposal consists of two phases:

- Phase 1 (present-end June 2015): This phase will involve the finalization of the design of the evaluations and their methodology; provision of technical assistance for design calibrations related to Evaluations 2, 3 and 4; the design of each Evaluation; roll out of Evaluation 2 in a limited sample; full roll out of Evaluation 3 and 4, and five rounds of tracker surveys. An evaluation report related to the *SFM* uptake evaluation (Evaluation 1) will also be produced during this phase.
- Phase 2 (July 2015-March 2017): This phase will consist of the roll out of Evaluation 2 on a larger sample to ensure adequate power to detect impacts; sixth and seventh rounds of tracker surveys that relate to all four evaluations as well as the end line surveys related to SFM, SFJ, and BPSV. The impact evaluation reports related to *SFM, SFJ*, *BPSV* and the synthesis evaluation report that summarizes the findings of all PEOP related evaluations will be produced during this phase.

The proposal is being conducted in two phases because DfID funding for the current program ends in June 2015 and hence surveys and outputs beyond this dates cannot be supported as part of the

[19] Expansion districts include Lahore, Gujranwala, Chiniot, Sargodha, Faisalabad, Narowal, Khanewal, Rahim Yar Khan, Sheikhupura and Vehari.

[20] These include Khanewal, Vehari and Rahim Yar Khan. Rahim Yar Khan borders the pilot districts of Bahawalpur and Muzaffargarh and Khanewal and Vehari border Lodhran and are located in the same division. Punjab's Multiple Cluster Indicator's Survey (2011) and Labour Force (2010-11) shows that labor market, literacy and education indicators in these three districts are very similar to the pilot districts that fall in the same divisions.

[21] The CERP PI's will also raise funds to support a part of the survey activity that equal one-third of the survey budget that is being submitted to DfID for support.

current program. The outputs and surveys committed under Phase 2 will be funded as part of the second phase of DfID's support for vocational skills training in Punjab, which is currently under consideration, provided it is approved. Separate deliverables and budgets of the two phases are detailed in section 4 of this proposal.

Data and Methodology

The evaluations will estimate the impact of the PSDF program on log frame outputs and outcomes where baseline values have been elicited through past surveys and on uptake among the target population. All evaluations will estimate the impact on a range of economic and non-economic outcomes that include: employment, individual earnings, household consumption, and an index of well-being and female empowerment, civic participation and state engagement. The baseline values of these outcomes have been collated for the treatment and control households as part of the Baseline Household survey as well as a short baseline outcome tracker conducted just before the interventions.

The purpose of each evaluation is to provide an estimate of cost-effectiveness and the value for money associated with each intervention. A synthesis document will be produced at the end of the evaluation phase that summarizes the findings and recommendations of all the evaluations and provides an analysis of relative cost-effectiveness and value for money associated with different components of the PSDF portfolio evaluated during this phase of the proposal. Given the timeline of the program, an estimate of the sustainability of impact of different interventions will be provided after one-year of their implementation[22]. We would be extremely keen to revisit the sample and evaluate sustainability of impact after three years but this timeline falls outside the timeframe of the current program and will add significantly to the budget.

For each evaluation these outcomes will be followed through multiple trackers and an end-line survey. Individual outcomes will be follows through periodic short tracking surveys, whereas household and individual outcomes will be tracked through end-line surveys. Measured individual outcomes will include the following individual economic outcomes: skills attainment (core and vocational); monthly earnings; labor force participation; occupational status[23]; employment status[24]; variability in annual earnings and employment status. In addition, these trackers will also obtain data that will allow non-economic outcomes of interest to be tracked through indices of well-being, physical health, mental health, state engagement and civic participation; female empowerment; male and female attitudes towards gender equality and women's attitude towards paid work. It will also field a set of questions about program leakage and non-compliance on behalf of providers, which will ask beneficiaries whether they received the services associated with specific interventions and on terms set by the program. The main source for eliciting household outcomes will be the end line surveys. These outcomes include: skills attainment of members; educational attainment of members; household income, consumption and savings; poverty and vulnerability level; food consumption; earnings of members; transfers and remittances in the household; labor force participation among

[22] The only exception is Evaluation 2 (Market linkages for rural women) where impact will be provided after six months of the intervention rollout.

[23] Occupational status classifies an individual's occupation based on the ISCO classifications of occupations.

[24] Employment status classifies individuals into whether they are paid employees (other than daily laborers), daily laborers, self-employed, apprentices, unemployed and seeking work, unemployed and not seeking work, engaged in unpaid non-household work or unpaid household work.

members; occupational status of members; employment status of members; depth of job search networks; asset ownership; health status of members and gender equality and empowerment indices. Impact will be measured on these individual and household level outcomes for all four evaluations.

The evaluations will continue to follow a methodological approach (which gave successful results in the baseline phase) that integrates quantitative and qualitative methods and includes the following components:

- Since causal inference is central to understanding the success or failure of different interventions, we opt for the randomized control trial (RCT) design, which is commonly regarded as the gold standard method for determining causality[25]
- Quantitative instruments will include baseline, end line and tracker panel surveys that enumerate household and individual outcomes (see above) in treatment and control households for the four different interventions that will be evaluated during this phase of the work.
- Qualitative analysis will be based on key informant interviews, open-ended interviews with small-N samples and focus groups that will be an integral part of each evaluation. Qualitative information gathered through these instruments will be designed in a way that: provides pre-implementation analysis on the potential of different intervention designs being proposed by PSDF and identifies factors that are likely to create and mitigate impact; documents the intervention process and compliance with different components of the intervention; provides a better understanding of different channels that underlie the causal impact of different components of the intervention identified through quantitative analysis; and analyzes the constraints imposed by non-intervention factors that may have reduced or offset the impact of each intervention. The qualitative analysis will be documented through structured field reports that provide a detailed analysis at the trainee, household, community, market and/or employers levels as is relevant for different interventions.
- Engagement with training providers and employers through surveys and qualitative fieldwork to obtain feedback on the design of interventions and quality of trainees respectively. The engagement with training providers in all evaluations is essentially built around the treatment monitoring compliance activity. We will use this activity to obtain valuable feedback for PSDF on the specifications of the different training scheme. Employer surveys that provide feedback on trainee quality and qualitative interviews and key respondent interviews with employers are an integral part of Evaluation 3 (*SFJ* with job placement), which evaluates PSDF's core training scheme that is focused on job seekers as opposed to the self-employed.

Structure of This Document

This proposal is organized as follows:

[25] See http://www.povertyactionlab.org/methodology

- Section 3 provides details about the four prospective evaluations proposed for the Evaluation Phase. Section 4 provides a detailed work plan and core outputs of the evaluation phase
- Section 5 describes the expected impact of the collaboration.
- Section 6 considers the wider impact and sustainability of the project by elaborating on the dissemination strategy and providing details about how the project will contribute to building local capacity for evidence-based policy design and impact evaluations.
- Section 7 describes how the Evaluation Phase builds on work already completed in the Baseline Phase of the project.
- Section 8 provides details of risk and risk mitigation strategies.
- Section 9 provides details of the CERP team who will work on this project

3. Details of the Evaluation Phase Proposal

This section provides details about the methodology, the evaluation design, the sample size, and the survey requirements of the four evaluations summarized in the previous section.

Evaluation 1: Skills for Markets (SFM) Scheme plus interventions to mitigate access constraints

Overview

The *Skills for Market* (*SFM*) scheme launched by PSDF targets marginalized populations with lower levels of formal education to open up opportunities for skills acquisition and increasing earnings by augmenting human capital. Rural women are an important focus of the scheme as more than 80% of rural women in the pilot districts have less than primary attainment and a large proportion of this group lives in acute poverty. A RCT based impact evaluation of the SFM scheme for rural women is being currently rolled out as part of PSDF's 2013-14 *Skills for Market* scheme (see Appendix C for progress on the *SFM* 2013-14 scheme). This evaluation has been designed to measure the impact of:

(a) Acquiring the popularly demanded training in tailoring on the economic and non-economic returns to rural women in the pilot districts

(b) Design calibrations to address salient access hurdles that adversely impact uptake and economic returns in a social context where women face severe mobility constraints.

Impact Evaluation for the SFM Scheme

The evaluation will measure impact on two distinct populations: (a) a representative sample of the general rural population of the pilot districts and (b) the selected population who chooses to self-enroll in the course. Impact on the former population will be measured on the basis of CERP sample households who are given the offer of training in the form of a voucher that gives them priority in admission.

Evaluating impact on both populations is of great policy relevance. One view is that the more meaningful population to consider in the evaluation of skills programs is the population that self-enrolls in courses (Attanasio 2009) because it is more interested and may possess superior ability to those who do not show up. However, in a context with low mobility it is extremely likely that able and willing individuals from poor households are unable to access training because of social and economic constraints. Moreover, the net value add on this population may be higher than that of the self-selected one, given that the latter may be motivated enough to find alternative avenues even if PSDF course options were not accessible. Therefore evaluating the impact on both populations is of great policy importance as it provides evidence on what types of populations to target and which populations offer higher returns. The *SFM* evaluation provides a unique opportunity to do this.

Evaluation Methodology for the SFM Scheme

The evaluation of the impact of the *SFM* scheme on the general population uses an encouragement design where vouchers for training and an offer of additional interventions to mitigate access

constraints (detailed below) are given to a set of randomly selected individuals (the treatment group) from the CERP baseline survey. Their outcomes will be compared to a randomly selected of individuals from the survey who are not provided vouchers and are not given an offer of the additional interventions (the control group). The randomization guarantees that on average these groups are identical on all factors that might influence enrollment and outcomes except for receiving the offer of training. Comparing them then gives an accurate estimate of the program's impact on various outcomes. Since individuals from the control group may also be able to acquire training[26], we will use an encouragement design where encouragement through vouchers will be used to instrument for actual receipt of training.

In addition, we will also be able to use an oversubscription design to evaluate the impact of the scheme on the selected population who applies for training.[27] The oversubscription is possible because there is high demand for the tailoring course among rural women, as indicated by the CERP baseline survey, and providers have been asked to randomize admissions in case of excess applications. Impact will be evaluated by comparing the outcomes of those women who are offered admission (treatment) to those who are not (control).

Design Calibrations to Mitigate Access Constraints for Rural Women

The baseline phase revealed that uptake of training among rural women is severely constrained due to access barriers that are correlated with the distance between a potential trainees' village of residence and the training centre. We began with a baseline assessment of households of the demand for skills training (a sample size of 31,495). Given evidence of strong demand, we conducted an RCT to measure the impact of skills training on economic and social outcomes as part of PSDF's *Skills for Employability (SFE)* 2011-12 scheme. Unfortunately, many women were unable to enroll despite large expressed demand: of 464 women offered training options 50% also took the additional step of selecting a specific course but only 5% were able to ultimately enroll. Results also showed that very few people from the sub-populations of greatest interest to PSDF, women belonging to poor and vulnerable households had relatively low uptake in the general population. These findings raised the concern that the sub-populations of interest to the program were not participating enough to benefit from the trainings supported by it. CERP conducted over 200 qualitative follow-up interviews and focus groups and a pilot RCT evaluation that addressed distance constraints through village-based training. This research revealed an interlocking web of constraints, obligations, norms and logistical difficulties that impeded enrollment and provided evidence on economic and non-economic factors constraining access. The results of the pilot RCT evaluation show that a meaningful proportion of the target population, which include women in general and those belonging to poor and vulnerable households, will remain excluded from training without meaningfully addressing these access constraints.

In response to this evidence PSDF has introduced design calibrations as part of *SFM* to address salient access hurdles and to examine the impact and relative cost effectiveness of these interventions on training uptake among rural women. These interventions have been designed and are in the process of being rolled out as part of the *SFM* 2013-14 scheme and are embedded in the

[26] The data from the *SFM* 2012-13 evaluation shows that non-compliance among controls is not an issue as less than 0.5 percent of control households enrolled in skills training.

[27] As opposed to the impact on general population which will be obtained through the encouragement design, as mentioned above.

roll out of the full evaluation of *SFM*. In particular, these interventions are designed to address the following access constraints:

1. Physical Distance: In many developing countries physical distance traveled by women is correlated with a number of interlocking constraints; by placing training centers in villages we can alleviate this constraint. Also, including the additional interventions (given below) in some villages that receive training centers can help determine which underlying factors contribute to the distance access constraint.

2. Safe and Reliable Transportation: While literature in multiple developing country contexts cite distance as a significant barrier to participation (Solotarof 2012; Mani 2012; Kabeer 2012), this constraint is not simply about geographic distance. Importantly, safety concerns that amount from unreliable or dangerous modes of transportation present additional barriers to access. This intervention will provide reliable group transport services to women trainees.

3. Lack of Information: While many women have expressed interest in skills training courses, they may have questions about the content and expected earning potential, which require credible sources of information. Lack of such credible information could be causing women to undervalue the courses. Additionally, an increasing amount of literature in behavioral economics suggests that nudges, such as reminders or invitations, could be enough to convince people to act on their preferences (Thaler and Sunstein 2008; Banerjee and Duflo 2011). The intervention here consists of meetings with trainees in which we share salient information on returns and provide opportunities to visit the training center with other female trainees from their village.

4. Financial and Credit Constraints: Even if they are free, courses involve opportunity costs, including forgoing household work and additional income earning potential. Providing varying stipends to trainees will enable us to quantify this effect.

5. Social Norms: Even if women want to attend courses, they may feel unable to do so because of restrictive social norms (Wigfield 2012). Crucially, men see transgressing restrictive gender norms as impacting their reputation directly (Jamali 2009), and may be unwilling to allow women of their household to participate, even if they see its value (Naqvi and Shahnaz 2002). In addition to meetings with trainees, the intervention will also consist of separate meetings for other (especially male) members of households and broader community members, providing opportunities to community members to visit the training center.

Another strand of literature recognizes that constraints are interrelated, and that interventions may be more effective when combined (Bandiera et al 2012). We will test this hypothesis as part of the *SFM 2013-14* evaluation by analyzing a bundle of the above interventions to determine their combined impact.

By using an RCT design (different villages will be randomly selected to receive one of the above interventions) we hope to establish which interventions increase enrollment most cost-effectively and therefore which factors are more important in constraining access. The results will provide extremely important information on gender access in the skills and labor market. In addition, the

evaluation will assess the socio-economic returns to training and the additional effect of linking to market opportunities. In addition to the RCT design, we have raised additional funds to support detailed focus groups with the help of experts in gender analysis in order to examine the more qualitative impact of these interventions and obtain a better sense of the underlying mechanisms at play.

Sample Size

The combined treatment and control group sample for the evaluation consists of 13,700 households in 371 villages across 3 districts of Southern Punjab. This includes a sample of 10,700 households from the general population and a sample of 3,000 from selected population. The general population sample comes from the CERP baseline survey which is drawn from, and representative of[28], the population in the PEOP pilot districts. The selected population sample comes from the list of self-selected applicants at each training centre. Power calculations suggest the need to have a larger general population sample to detect effect sizes of interest to the log frame. This is because uptake is relatively low in this population, which is likely to weaken the first-stage estimates. The higher uptake among the selected population suggests the need for a smaller sample to detect effect sizes of interest.

Households in this evaluation sample will be tracked through 5 rounds of tracker surveys of trainees and an end-line household survey. The first and second rounds of trackers will allow us to estimate the impact of design calibrations to mitigate access constraints on uptake and retention among the general population. All 5 rounds of trackers and the *SFM* end line survey will allow us to measure impact on economic and non-economic outcomes. Multiple rounds of tracker surveys are needed to gain power to detect impact on outcomes of interest; this is also based on power calculations conducted as part of the design and is supported by the literature on randomized evaluations (Mckenzie 2011). Power calculations suggest that a minimum of 3 rounds of trackers may just allow us to detect the effect size for different outcomes specified in the PEOP log frame. Therefore, we will be budgeting for only 3 rounds of trackers in this proposal. However, in order to detect the effect size specified in the log frame with certainty we will raise additional funding for 2 rounds of trackers.

Multiple rounds of data collection are not only beneficial to gain power to detect impact on outcomes, it will also help identify how outcomes change over time, and specifically, whether gains from skills hold stable, level off or depreciate over time. The trackers and end line surveys together will allow us to measure sustainability of program's impact, on several individual and household-level outcomes of interest, after one-year of the completion of *SFM* training.

Quantitative surveys are being supplemented with qualitative focus groups, key informants interviews and open-ended interviews with a small-N sample. These instruments have been used to capture detailed qualitative data on the potential for impact of different interventions to increase access and to document in detail factors that are likely to constrain or catalyze the impact of these interventions. In the case of the SFM intervention, the emphasis of this work was on documenting

[28] The sampling of these households was done using two-stage random sampling with a probability-proportional-to-size (PPS) sampling of villages and urban blocks in the first-stage (conducted by the Pakistan Bureau of Statistics) and a simple random sampling of households in the second stage (conducted by CERP after a complete listing of households in the sample villages/urban blocks).

constraints and the feasibility of interventions at the level of the household and the community. This includes access interventions documented in this proposal as well as others that had shown up in the survey exercise. This information has been used to provide evidence on design to PSDF.

Qualitative fieldwork has also been used to document the robustness of the implementation process and compliance with different components of the interventions including the provision of training itself. It has also been used to understand why certain trainees choose to enroll while others do not. Focus groups and key informant surveys have been conducted in a random sample of 10 treatment villages stratified by whether they got a village-based training centre and the distance of a village from the training centre. Key informant interviews were conducted with village influentials; teachers; women and men working as tailors; male and female activists and mobilizers; and male and female heads of families. Open-ended interviews have been conducted with a random sample of around 100 voucher holders in these villages stratified by whether they were given a voucher for training and those who accepted and rejected an offered voucher. We plan to repeat this qualitative field exercise after the completion of training to obtain a better understanding of different channels that underlie the causal impact of different components of the intervention and also document the non-intervention factors that affected the impact of the interventions. We will conduct qualitative work with treated and untreated individuals to understand their subsequent labor market experiences and this may suggest actionable constraints that PSDF may be able to address. Finally, qualitative interviews will be conducted with trainers to monitor treatment compliance and to obtain feedback for PSDF on the specifications and sustainability of the *SFM* scheme.

Evaluation 2: Market linkages for rural women

Overview

Evidence from the baseline household and employers surveys shows that women are predominantly engaged in home based work because of the reluctance of local employers to hire women and due to low labor market mobility. The social context that restricts labor mobility for women also reduces access to markets. This results in low economic returns to home-based work. There is recognition in the PSDF Board that economic returns to women's training cannot be realized without designing an effective post-training market linkage component. The intervention will develop a linking mechanism by matchmaking between existing or potential middle men who are linked to the demand side, and the pool of PSDF trained women. These middle men or sales agents will link women trainees directly to dense demand clusters in urban and rural markets that can both provide inputs to the women and lower production risks.

The design calibration will be implemented with an initial sample of women who successfully complete the *SFM* 2013-14 training. The sample for this evaluation will be increased in phase 2 of the collaboration with DFID to ensure that there is enough power to detect results. The evaluation is being split into two rounds for the following reasons:

1. Funding Constraints: Keeping in mind that the collaboration with DFID is to be broken into two phases, we have to ensure that the total cost for phase 1 satisfy the constraints.
2. Adoption of Phased Evaluation Approach: The two-round design is being proposed because of the success of the phased evaluation approach in the baseline AG, which evaluates the

direction of impact on a small sample in the first round to attain objective evidence for scale-up in the second round.

This evaluation will measure the additional impact of strengthening market linkages on economic and non-economic returns to training. Seeding an effective model of training that integrates market linkages is of great policy importance as it will help address concerns about public investment in skills training for women on the grounds of low economic returns.

Evaluation Methodology for the Market-Linkage Intervention

The evaluation of market linkages will use an RCT design where a post-training market linkage intervention will be provided for several months to graduates (treatment group) in a random set of *SFM* (2013-14) training centers. Their returns will be compared with the returns of graduates from the remaining (control group) *SFM* (2013-14) training centers.

Sample Size

As mentioned above, the market linkage intervention will be done in two phases. In the first phase an initial sample of eighteen *SFM* (2013-14) training center villages will be chosen. This will be followed by an expanded sample of thirty two SFM (2013-14) training center villages in Phase 2 of the evaluation phase. The expectation is that approximately half of the graduates from a center or 10 trainees per center will avail this intervention. Therefore the treatment group will consist of 500 graduates who will be provided market linkages (180 graduates in Phase 1; and 320 graduates in Phase 2). We plan to contain costs by exploiting synergies between the *SFM* and market-linkage evaluations and will use the 3 rounds of *SFM* trackers to track outcomes for this evaluation as well. The trackers and end line surveys will allow us to measure sustainability of impact after six months of the roll out of the market linkage component.

In addition, we will create synergies with the qualitative fieldwork being conducted for Evaluation 1 to assess the feedback on the market linkage design being proposed by providers to PSDF as well as check the underlying assumptions of this design. Particular attention would be paid to analyze constraints and opportunities at the level of households, the community and related to market access and mobility. We plan to use phone interviews with a small-N sample to assess the robustness of the implementation process and compliance with the proposed design and supplement the phone interviews with field visits to households in areas where compliance violations are being identified. Finally, qualitative instruments will be used at the end of the intervention period to understand the channels through which impact has been created and also to document the role of important non-intervention factors. This component of the fieldwork will be conducted with treated and untreated individuals on their subsequent labor market experiences as well as those that take up market linkages versus those that do not. Finally, qualitative interviews will be conducted with trainers to monitor treatment compliance and to obtain feedback for PSDF on the specifications and sustainability of the market linkage scheme.

Evaluation 3: Skills for Jobs (SFJ) Interventions plus calibrations that integrate job matching

Overview

The *Skills for Jobs (SFJ)* program offers training courses in a variety of technical fields, followed by testing and certification by the Punjab Board of Technical Education (PBTE). The redesigned PSDF program has added increasing productivity in sectors with growth and employability potential as an additional objective for the expansion districts. This has led to the introduction of skills schemes in the *SFJ* mode in sectors with growth and employability potential in the expansion districts.[29] *SFJ* in the pilot and expansion districts and these sectoral variants in the *SFJ* mode comprise the major share of PSDF's training portfolio – out of the 38,000 trainees that have been trained by PSDF to-date, 18,500 (almost 50%) are from the *SFJ* scheme.

The evaluation of *SFJ* is designed to measure the impact of:

(a) The *SFJ* scheme (including its sectoral variants) on economic and non-economic returns to trainees. CERP will conduct RCT-based evaluations of a sample of training courses being offered as part of the *SFJ* scheme and its sectoral variants. These evaluations will be conducted in a sample of pilot and also expansion districts.
(b) Design calibrations that integrate job matching services in these schemes on uptake and economic and non-economic returns.

The impact of the scheme and the additional impact of the job matching design calibration on those who attend training will be evaluated for two populations:

(a) A representative sample of the general urban population of the pilot districts and
(b) The selected population who chooses to self-enroll in the relevant courses to be evaluated in the pilot and the expansion districts.

Impact on the former population will be measured on the basis of CERP sample households in urban areas of the pilot districts that have been surveyed during the baseline phase. The impact on the selected population will be well powered and able to detect reasonably sized outcome effects. The statistical power for the general population is likely to be lower because of low uptake and it may only be able to detect larger outcome effects. However, general population estimates are critical for understanding the potential overall impact of the program. Doing both evaluations also allows us to compare differential selection effects. This is really important information for program targeting. The evaluation of the general population will allow us to analyze the value individuals in the general population place on job matching services by estimating differential uptake among those who are given a voucher for admission and those who are given a voucher along with job matching services. By evaluating the impact of the program on the selected population who self-enroll we can benchmark it's value-added for trainees. Although the results of the evaluations in pilot and expansion districts may not be strictly comparable the advantage of doing this evaluation in the urban areas of both types of districts is that it will allow us to check whether there is potential heterogeneity in impact across urban environments, which is valuable information for PSDF.

[29] The sectors in which skills schemes are being introduced in the *SFJ* mode are: Garments, Food Processing, Leather and Shoe Making, Light Engineering, and Logistics.

Design Calibrations that Integrate Job Matching Services

International evidence shows positive and significant returns to vocational training in developing countries in skills programs that pair in-class training with job placement/access interventions (Attanasio et. al. 2009, Hicks et. al. 2013, Card et. al. 2011, Nopo 2007). Evidence from the baseline phase shows that job placement and search networks are exclusive and highly personalized, which is likely to adversely affect the probability of getting employment and, in turn, lower uptake and returns. The baseline survey also shows that there is a stated desire among both job-seekers and employers for a more standardized process to match laborers with appropriate opportunities. There is evidence from the pilot SFJ evaluation conducted in the baseline phase[30] that the uptake of *SFJ* training has been low (below five percent) in the pilot districts among urban males who are actively participating in the labor market and is unresponsive to doubling and trebling stipend levels. Qualitative follow up interviews done as part of the *SFJ* pilot evaluation also reveal that individuals who are actively participating in the labor market do not want to give up their current job to gain better skills because they are unsure about the probability of getting jobs after completing their training. Therefore, integrating job matching interventions with the flagship training programs is an important design calibration that promises high returns. Seeing this potential the PSDF Board has approved the design of a pilot that integrates job matching services and has requested CERP to conduct a RCT-evaluation to estimate the returns associated with this intervention for the urban population in the pilot districts and for a sample of expansion districts.

The design of the job matching intervention will build on the current efforts of PSDF to build a network of registered local enterprises and firms that act as a pool of potential employers in the pilot and expansion districts. The job matching intervention will augment the current trainee profiles with report cards prepared by TSP that are designed to provide employers information about a trainees technical/cognitive ability; non-cognitive ability (such as professionalism, reliability, integrity and interest in acquiring a job) and past job experience. The exact dimensions of these report cards will be designed in consultation with PSDF. The job matching intervention is predicated on the hypothesis that search costs constrain both employers and employees and so long as training is of the right type and of good quality we expected a positive response from employers to trained skilled workers given that PSDF's employment registration drive will have built a credible network with employers. We will, on DfID's suggestion, use a employer feedback survey to ask whether existing job search networks act as substitutes or complements for job matching. The matching intervention will use PSDF's reputation as the premier institution of skills training to intensively provide this information to employers with the aim of catalyzing job placement. The design details will be finalized by designated members of the PSDF Board and Management with evidence-based technical assistance from CERP.

Evaluation Methodology for the SFJ Scheme

The *SFJ* impact evaluations will use a combination of an encouragement RCT design for the general population and an oversubscription RCT design for the selected population, which is consistent with the more rigorous vocational training evaluations in developing countries (Attanasio et. al. 2009, Hicks et. al. 2011, Card et. al. 2011). In the expansion districts the impact of the schemes and the additional impact of the job matching design calibration will be evaluated for the selected population who chooses to self-enroll in the relevant courses to be evaluated. By evaluating the impact of the

[30] Cheema et. al. (2013b)

program on the selected population we can benchmark the value-added of this training for trainees self-enrolling in courses. We will be unable to use an encouragement RCT design to measure the impact of this training on the general population in the expansion districts because of the absence of a baseline in those districts. Given the extension of the program to the expansion districts at this late stage it would not be cost or time effective to construct a baseline for purposes of intervention design, targeting and evaluation of the general population in the expansion districts.

The encouragement design will consist of offering training vouchers to a randomly selected subset of the pilot households (treatment group). The oversubscription design will work as follows: each training provider participating in the evaluation will be asked to provide a list of suitable applicants that is 50% longer than their contracted capacity and these applicants will be randomly offered (treatment group) or not offered (control group) admission in the training institution. Those who are not offered admission in the evaluation round may be offered deferred admission in a future round. An application information system will be set up to register applicants as part of the evaluation and its deployment will be outsourced with technical assistance from CERP and oversight on this will be provided by PSDF. An advantage of this design is that random assignment allows us to overcome selection bias that casts doubts on the validity of the estimates, which could be biased if there is selection into the program on the basis of unobservable characteristics. The impact on the selected population will be well powered and able to detect reasonably sized outcome effects. The impact of job matching services will be measured by an additional RCT that randomizes the intensity of access to job matching services.

Sample Size

The sample of the *SFJ* evaluation consists of 7,500 households divided into treatment (two-thirds of the sample) and control groups. This will include a sample of 4,000 individuals from the selected population and a sample of 3,500 CERP baseline households that is representative of the general urban population of the pilot districts. The sample of the selected population will be drawn from the pool of applicants to *SFJ* training courses (in high demand) and will be representative of the selective population who are motivated to self-enroll in these courses. The general population sample comes from the CERP baseline survey which is drawn from, and representative of[31], the population in the PEOP pilot districts. Two-thirds of the treatment sample will be randomly provided the additional job matching service after training. Power calculations show that the size of the selected population sample is well powered to be able detect reasonably sized outcome effects. This sample size is comparable to the size used in other studies that use an oversubscription methodology (Attanasio et. al. 2009, Hicks et. al. 2011, Card et. al. 2011). The statistical power for the general population is likely to be lower because of low uptake and it may only be able to detect larger outcome effects.

The sample of the *SFJ* evaluation (and its sectoral variants) in the expansion districts will consist of up to 3,000 households divided into treatment and control groups. Baseline data for the treatment and control samples in the expansion districts will be collected through baseline tracker surveys that precede training. This will be followed by two further rounds of post-training tracker surveys. End

[31] The sampling of these households was done using two-stage random sampling with a probability-proportional-to-size (PPS) sampling of villages and urban blocks in the first-stage (conducted by the Pakistan Bureau of Statistics) and a simple random sampling of households in the second stage (conducted by CERP after a complete listing of households in the sample villages/urban blocks).

line surveys will not be conducted with the treatment or control groups in the expansion districts. The same 3 rounds of trackers will be used to measure the impact of *SFJ* training (and its sectoral variants) as well as the impact of the job matching. In addition, based on feedback from DfID we will conduct 3 rounds of tracker surveys with a sample of between 500-750 employers of PSDF graduates, in the expansion districts, to gauge employer feedback and satisfaction about the graduates that they have hired. Multiple rounds of surveys will also be useful to identify how outcomes change over time and whether gains from skills sustain, level off or depreciate over time, which is really important information for PSDF and other policy makers. The post-training tracker surveys will allow us to measure sustainability of impact after one-year of the completion of *SFJ* training.

The oversubscription RCT with the selected population will be done for a random sample of *SFJ* courses that are in high demand in the expansion and pilot districts and will provide an impact of the more popular PSDF courses as a whole. The exact sample of courses will be finalized in consultation with PSDF and DfID during the design phase. The mapping of high demand courses will be extremely valuable information for PSDF in itself and it will also determine the sample of expansion districts in which the evaluation will be conducted.

Households in this evaluation sample will be tracked through 4 rounds of tracker surveys of trainees as the added rounds of data collection are necessary to gain power to detect impact on outcomes of interest. Round 4 of the trackers will precede training and provide an initial baseline for the selected population sample in the expansion districts and round 5 of the trackers will provide an updated baseline for the general population sample. The same 4 rounds of trackers will be used to measure the impact of *SFJ* training as well as the impact of the job matching. In addition, based on feedback from DfID we will conduct 4 rounds of tracker surveys with a sample of up to 1,000 employers of PSDF graduates to gauge employer feedback and satisfaction about the graduates that they have hired. Again, multiple rounds of surveys will also be useful to identify how outcomes change over time and whether gains from skills sustain, level off or depreciate over time, which is really important information for PSDF and other policy makers. The trackers and end line surveys will allow us to measure sustainability of impact after one-year of the completion of *SFM* training.

Quantitative surveys will be supplemented with key informants interviews and open-ended interviews with a small-N sample. An important purpose of the *SFJ* qualitative field survey exercise will be to document the type of cognitive and non-cognitive information that trainers can potentially generate and the type of trainee-level information that matters to employers of different types and scales while making their hiring decisions. This evidence will be provided to PSDF for the purposes of design of the job placement intervention. This information will be collated through key informant interviews and open-ended interviews with a small-N sample of employers registered with PSDF and training providers running courses that are over-subscribed. This exercise will also be used to document the factors that employers take into account while making the employment decision and what type of information and other support would they require to seriously consider applicants that do not come through the traditional job search networks. The sampling strategy and the sample size for this work will be decided after consultation with PSDF and after documenting the degree of oversubscription by training provider and course type. We will also conduct open-ended interviews with existing PSDF graduates and potential trainees to develop an understanding of how to market the job placement intervention to prospective trainees; what type of information increases their expectation of getting a job after training and the constraints to getting better jobs. The sample for this activity will be stratified by: whether the respondent obtained PSDF training or not; whether he

chose to complete training after enrollment or not; and by labor market status. This qualitative survey exercise will also yield an understanding of the different channels underlying the estimated impact. Qualitative interviews will also be conducted with trainers to monitor treatment compliance and to obtain feedback for PSDF on the specifications and sustainability of the *SFJ* and job placement interventions.

Evaluation 4: Big Push on Skills for Villages (BPSV) Scheme

Overview

Agriculture, livestock and related sectors are the biggest employers of labor and the also see the greatest demand for skills in the pilot districts (Cheema et al 2012). However, output growth rates in these sectors have been sluggish reflecting declining total factor productivity growth rates (Ahmed and Gautam 2013). The Planning Commission (2009) finds considerable yield gaps between the average agricultural producer in these sectors and progressive farmers. Similarly, the CERP baseline report on livestock and dairy finds enormous variation in milk yields per animal – a measure of productivity in the livestock sector, with the 25% most productive households having productivity levels more than double those of the least 25% productive households (Rasul et al 2012). These productivity differences remain even after accounting for some basic differences across the two types of households such as the number and composition of livestock owned, livestock breed, and characteristics of the household head. The report finds that practices and inputs used towards livestock differ across these groups, and recommends the need for interventions that provide information on best practices and make available basic inputs and veterinary services (Rasul et al 2012). The PSDF Board is of the view that increasing productivity through training in best practice skills has the potential to increase productivity in these high employability sectors, which evidence suggests promises high benefits in terms of poverty reduction (Malik 2005). The CERP baseline survey report also provides evidence of large-scale demand among the village population of the pilot districts for training in these sectors. In the view of the PSDF Board, together these are essential pre-conditions for inclusive growth and the successful implementation of this program has the potential of high impact. In spite of this potential, the provision of these skills constituted only 11% of the current stock of PSDF graduates because of the paucity of finding cost-effective off-the-shelf providers who can supply these skills.

Recognizing this potential the PSDF Board has approved the design concept of the *Big Push on Skills for Villages* (*BPSV)* scheme, which aims to increase the productivity of rural agricultural communities through saturating villages with intensive training in agriculture, livestock, and veterinary best practices throughout the value chains. The design will aim to shift the skills and practice frontier (and hence the human capital frontier) at the *village*-level and create large-scale impact by exploiting both scale economies and production complementarities. The objective is to identify whether saturating villages with human capital in high employability sectors that will catalyze scale economies and complementarities can act as an engine of growth. This is very much in line with ideas like Sach's Millennium Villages and consistent with "big push" theories of economic growth. Identifying a role for skills training in bridging yield gaps builds on a substantial foundation of research that finds that agricultural extension and training can be successful in eliminating some of the constraints on agricultural productivity in low-income settings (IFC 2013). The extent to which this finding holds, however, depends on the context and the package of interventions, as well as their method of delivery (Hanna 2012, MCC 2012, BenYishay 2013).

The scheme, which will be finalized during the design phase, will be based on a review of the local and international evidence on the success of different designs of training and technical advisory (TA) service models for villages and through consultations with stake holders as well as existing and potential providers of training and related services. The design involves PSDF developing a comprehensive menu of trainings in frontier skills and practices related to the entire value chain in these sectors. The design of this menu will involve an engagement with a consortium of leading private sector agricultural companies and input suppliers as well as progressive farmers to share knowledge on frontier skills and practices across the value chain. The PSDF Board is proposing to build this consortium through a corporate social responsibility initiative spearheaded by members of the Board with depth of networks and experience in these sectors. The delivery of this menu will rely on the standard PSDF procurement model that has been extremely successful in seeding new suppliers of skills in different sectors. This design will also involve a Training of Trainers component for the trainers of the selected TSPs.

The design phase will be spearheaded by designated members of the PSDF Board and Management. In addition, to support the design efforts of the Board we propose to engage a specialist in the area of design for rural farm and non-farm training. The design phase will precede the evaluation phase. CERP's main role is to test and provide evidence on the evidence associated with this development approach. CERP is agnostic about the direction and scale of the impact *apriori*. There are good theoretical reasons for this approach to work and if it does it could promise large impact. However, CERP also recognizes that there could be other theoretical reasons to expect low impact from this approach such as risk aversion in adopting new practices. CERP is not advocating this approach but aims to rigorously evaluate its impact potential.

Evaluation Methodology for the BPSV Scheme

The *BPSV* scheme will be evaluated using 90 villages from the in-depth sample of villages. Impact will be evaluated by comparing outcomes in a random set of in-depth villages in which this scheme is introduced (treatment villages) to the remaining set of in-depth villages in which this scheme is not introduced (control).

The existence of a large household sample[32] in the baseline in-depth villages, which allows the estimation of village-level measures (such as income, wealth, human capital stock etc.) has created the novel opportunity to rigorously measure the impact of human capital infusion through skills training on village level GDP and productivity, gains to households at different points in the income distribution, and the extent of spillovers that result from interventions. The availability of this baseline sample offers the potential to measure these outcomes at the level of the household and the village economy. Thus we will be able to estimate the impact of such "big push" interventions at the individual level for both direct (trainees and their households) and indirect beneficiaries (households who are not trained but live in the village that receives training) and ultimately, at the community/village level. To our knowledge this will be the first ever RCT rigorous evaluation of globally of these increasingly popular "big push" style interventions and therefore will not only inform PSDF and skills development in Pakistan, but will more broadly offer insights into the cost effectiveness of such programs globally.

Sample Size

[32] Approximately 35% of the in-depth villages have been sampled as a baseline.

The sample for this evaluation consists of 12,700 households in 90 villages evenly divided into treatment and control groups. Power calculations show that for the given sample, two trackers and an end-line survey will be sufficient to detect impacts on economic returns. CERP is committed to raising funding that covers half the cost of the end-line survey and the cost of one of the two trackers. The trackers and end line surveys will allow us to measure sustainability of impact after one-year of the completion of *SFM* training.

Quantitative analysis will be supplemented with qualitative focus groups, key informants interviews and open-ended interviews with a small-N sample. An important purpose of the qualitative work will be to document the variations in best practices in different sample communities and to understand the factors and constraints creating this variation within a local area. Another purpose would be to document the bundle of support required to increase the uptake of best practice skills and understand potential channels of spillovers. The qualitative fieldwork will also be used to understand the structure of training provision that is best suited to provide the widest possible diffusion of best practice skills and the costs associated with different implementation strategies. This information will be used to provide design feedback to PSDF. The information will be collated on the basis of focus groups in 10-15 villages and key informant interviews with large, medium and small farmers and livestock owners and operators, input suppliers, buyers and progressive farmers and livestock owners and operators in the area. We will supplement key informant interviews with open-ended interviews with a small sample of farmers and livestock operators of various types. The qualitative exercise will also be used to subsequently document compliance with the implementation process and to obtain an understanding of channels underlying the estimated impact. Finally, qualitative interviews will be conducted with trainers to monitor treatment compliance and to obtain feedback for PSDF on the parameters and sustainability of the training scheme.

Summary of the Evaluation Phase Proposal

Table **1** gives a summary of the evaluations that will be conducted as part of the Evaluation Phase.

Table 1: Summary of the Evaluation Phase

RCT-based Evaluation	Measures Impact of:	Outcomes	Unit of Evaluation	Population of Interest
SFM (2013-14) for rural women plus interventions to mitigate access constraints **(Evaluation 1)**	Training in the high demand tailoring skills	Economic and non-economic outcomes	Individual and household	- Representative rural female population in Pilot districts -Female applicants self-enrolling in training
	Design calibrations to mitigate access constraints	Uptake	Individual and household	- Representative rural female population in Pilot districts
Post-Training Market Linkages for Rural Women **(Evaluation 2)**	Market linkages	Economic and non-economic outcomes	Individual and household	- Representative rural female population in Pilot districts
SFJ (and its sectoral variants) for urban population with and without job matching **(Evaluation 3)**	Training in course that are part of these schemes	Economic and non-economic outcomes	Individual, household and employers	-Applicants self-enrolling in these courses
	Pairing job matching with training in *SFJ* related courses	Economic and non-economic outcomes Uptake	Individual, household and employers	-Applicants self-enrolling in these courses -Representative urban male population in Pilot districts
BPSV scheme for the rural population **(Evaluation 4)**	Village-based training in skills related to agriculture, livestock and related sectors	Economic and non-economic outcomes Uptake	Village and household	-Representative village communities of the pilot districts and village populations engaged in these sectors in pilot districts

4. Phase-wise Work Plan of Interventions, Surveys and Outputs

This section provides the work plan for the impact evaluations that will be conducted during this phase of the AG agreement. It also provides the details for the data collection and surveys that will be done to collect individual and household level outcomes over time. Since the evaluations are being proposed in two phases (Phase 1 (present-June 2015) and Phase 2 (July 2015-March 2017), we provide separate details of the work plan, interventions, surveys and outputs for the two phases. Separate budgets for the two phases are being submitted as a separate file.

Phase 1 (present-end June 2015) Work Plan

This sub-section relates to the interventions, data collection activities and outputs to be delivered under the Phase 1 of the Evaluation Period i.e. from present to end June 2015.

Interventions in Phase 1

Evaluation 1: *SFM Evaluation for rural women plus design calibrations to mitigate access constraints*

Interventions to mitigate access constraints for rural women will be rolled out as part of the *SFM* (2013-14) scheme. The interventions to mitigate access constraints that will need to be funded through the TA include:

(a) Stipend top-ups (over and above the base stipend provided by PSDF) to analyze the impact of financing to offset the high opportunity cost of training
(b) Secure group transport to offset costs related to distance constraints.

PSDF has asked CERP to fund these interventions because financing these interventions through PSDF imposes significant operational costs as they will have to set-up an elaborate open bidding process which will be highly costly given that these interventions are being piloted on a small-scale. The cost of the SFM training itself is fully supported by PSDF and in addition a base stipend of Rs. 1,500 per trainee per month is also supported by them.

The proposed stipend top-ups will be randomly provided to a sample of 1,500 enrollees from the CERP voucher holding households upon successful completion of the course and the average stipend top-up is Rs. 2,250 per month. This amount has been calibrated to equal the monthly cost of transport plus forgone earnings for women engaged in unskilled farm labor.

Secure and reliable transport will be randomly provided to 260 enrollees from CERP voucher holdings who have agreed to avail it as a group and have regular attendance. The average monthly cost of this transport is Rs. 5,000 per trainee per month.

In addition, we will need a budget to monitor training provider compliance with the protocols of the interventions being rolled out to mitigate access constraints that include community mobilization, open days and admissions through ballot to ensure randomization.

Evaluation 2: *Market-linkage evaluation with rural women (small sample)*

PSDF has asked CERP to fund the costs of this pilot from the TA in order to reduce overheads that would result if the intervention was supported through PSDF funding. If proven successful in a benefit-cost sense, PSDF could then incorporate these costs into the additional rollout of such programs. The costs of market linkages in this evaluation phase will include time-bound support to SFM graduates and the costs of seeding market agents and their market connectivity. The aim is to wean the graduates of this financial support in a period of several months (likely 5 to 6) and ensure that they transition to a self-financed model in this period. The costs of production circles in this period includes: (a) the rental of building where training circles will be organized; (b) rental of sewing machines; (c) liquidity for raw material for a two-three month period; (d) base salary and commission for sales agents (e) marketing costs of their first two production cycles and (e) transport costs for marketing visits.

The plan is to organize market linkages for the graduates of a set of 18 random centers after *SFM* (2013-14) training has been successfully completed. The expectation is that 10 trainees per centre (or 50% of the class) will avail this facility after successful completion of training.

We have received proposals for the market linkages intervention from 5 providers. One of the five providers has experience with successful implementation of market linkages in the Pakistani context and another has well-developed networks with branded retail outlets such as Bareeze, Care Crafts and Leisure Club that are interested in carrying a dedicated line of products that are produced by *SFM* graduates. We expect to roll out the first market linkage component between January 2015 and May 2015 and the second component from August 2015 to January 2016. The choice of which model to employ will be finalized in collaboration with PSDF and after inputs from an expert hired for this purpose. The budget for the expert will be supported by CERP.

Evaluation 3: *SFJ Evaluation for the urban population plus job matching intervention*

We expect the design of the *SFJ* job matching intervention to be completed between September 2014 and January 2015. The timeline for the training and evaluation of *SFJ* will be staggered for the selected and general populations. The timeline for the training will be between April – November 2015. The gap between the finalization of design and start of training is to allow sufficient time for procurement of services for the scheme. The job matching intervention will be implemented between March-December 2015. This will include the production of report cards by TSPs and the dissemination of these report cards to registered employers.

The operationalization of the oversubscription design will require setting up an online application registration system that can be used by TSPs to register all applications. The deployment of this system will be outsourced with technical assistance from CERP and its oversight will be provided by PSDF. This system will be verified through the rigorous third-party monitoring system that PSDF has put in place. We have successfully deployed an oversubscription design with six training service providers in three pilot districts as part of *SFM* 2013-14 and are confident about a successful roll-out of this intervention. The application data will be used for randomization of admissions and is central to the roll out of the evaluation. This will need to be funded from TA.

In addition, the *SFJ* evaluation of the general population will require delivery of vouchers and information to 2,000 treatment households. The job matching intervention will require printing of

reports and delivery to a treatment sample of 5,000 households (2,000 from the general population and 3,000 from the selected population). In addition, we will need a budget to monitor training provider and household compliance with the protocols of the *SFJ* evaluation and its job matching component.

Evaluation 4: *BPSV Evaluation for the rural population*

The following timelines will apply to the roll out of this intervention:

- November 2014 – January 2015: Creating the consortium of best practice skills providers and getting a commitment from them on sharing their best practices.
- December 2014 – March 2015: Finalizing the design of the training of trainers and training components.
- December 2014 – March 2015: Completing the procurement cycle.
- June – August 2015: Training of trainer's component will be completed during this time. PSDF requires four months to complete procurement after the finalization of the design.
- August – December 2015: Delivery of *BPSV* training.

We will require a budget to support the delivery costs of information, vouchers and visits to monitor compliance with the interventions.

Tracker Surveys in Phase 1

Individual outcomes will be measured through seven tracker surveys related to the *SFM, SFJ and BPSV* evaluations; five of which will be conducted in Phase 1 of the evaluation proposal. The purpose of these trackers is to provide evidence of how outcomes hold-up over time as well as enhance statistical precision in the quantitative analysis for end of programme evaluation. The trackers have been designed to provide an updated baseline (for a reduced set of individual outcomes) and short-run and medium-term information for outcomes of interest for the program. The description, proposed time and sample size for each tracker is given below.

- Tracker 1 will be conducted in April 2014 with a sample of 13,700 households. This will measure the pre-training outcomes for the individuals in the *SFM* treatment and control samples.

- Tracker 2 will be conducted in July 2014 with a sample of 13,700 households. This will measure the immediate post-training outcomes for the individuals in the *SFM* treatment and control samples. The need for this additional tracker is suggested by power calculations.

- Tracker 3 will be conducted in November 2014 with a sample of 13,700 households. This will track post-training outcomes for the *SFM* treatment and control samples.

- Tracker 4 will be conducted in February – March 2015 with a sample of 16,700 households and up to 500 *SFJ* employers. This will provide updated baselines for 12,700 BPSV sample households and 4,000 households in the selected *SFJ* treatment and control sample.

- Tracker 5 will be conducted in April-June 2015 with a sample of 17,200 households. This will track post-training outcomes for 13,700 individuals in the *SFM* treatment and control samples and provide updated baseline for 3,500 households in *SFJ's* general population sample.

CERP will raise funding for 1 out of the 5 rounds of the trackers with 13,700 *SFM* treatment and control sample. This equals PKR 17.96 million which is 25% of the budget that is being requested from DfID to support the tracker survey activity in Phase 1.

Various protocols are set by CERP to ensure that data retrieved through surveys meets quality standards. These protocols have been included and explained in Appendix E of the proposal.

Qualitative Fieldwork in Phase 1

Qualitative fieldwork associated with each evaluation (detailed in section 3) will be conducted in three different stages: (a) prior to the finalization of the design of the intervention by PSDF; (b) during the period the intervention is being implemented; and (c) at the time of the first post-intervention tracker. We have raised funds to support the qualitative fieldwork associated with Evaluations 1-3 and are in the process of raising funds for the qualitative work related to evaluation 4 (Big push on skills for villages).

Table 2: Work Plan for Phase 1

Work Plan for Phase 1 (Current-end June 2015)	2014												2015												2016
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1
INTERVENTIONS																									
Evaluation 1: *SFM Evaluation plus interventions to mitigate access constraints*																									
a) Design of Evaluations - (*complete*)																									
b) Design Calibration of Scheme																									
i. Design of village base training; transport; mobilization; and stipend interventions – (*complete*)																									
c) Roll-Out																									
i. Training under SFM scheme																									
ii. Provision of Transport																									
iii. Delivery of Stipend																									
Evaluation 2: *Post-training market linkages for rural women*																									
a) Design of market linkage intervention																									
b) Roll-Out of market linkage intervention 1																									
Evaluation 3: *SFJ (and its sectoral variants) Evaluation plus job matching*																									
a) Design of job matching intervention																									
b) Roll-Out of Intervention																									
i. Training under SFJ scheme (selected population)																									
ii. Training under SFJ scheme (general population)																									
iii. Job Matching Intervention																									
Evaluation 4: *BPSV Evaluation*																									
a) Design of Evaluation and Scheme																									
b) Roll-Out of Intervention																									
i. Training of Trainers																									

ii) Training under BPSV scheme																									
SURVEYS																									
a) Tracker Surveys																									
i. Tracker Survey 1 (HH sample = 13700)																									
ii. Tracker Survey 2 (HH sample = 13700)																									
iii. Tracker Survey 3 (HH sample = 13700)																									
iv. Tracker Survey 4 (HH sample = 16700; Employers sample = 500)																									
v. Tracker Survey 5 (HH sample = 17200)																									
Reports																									
a) SFM-B Uptake Report																									

Expected Outputs and Dissemination Activities in Phase 1

Findings from the evaluations will be provided in the form of notes and reports that are listed in Table 4. Dissemination activities are listed in Table 5.

Table 3: Publications by CERP for Evaluation Phase 1

Publications	Detail	Expected Timelines
Report on Design and Compliance of Uptake Interventions in SFM 2013-14	This report will detail the motivation, objectives, design of interventions, and evaluation design of calliberations to PSDF's SFM 2013-14 scheme aimed at increasing access to training for rural women. It will also report on the sample for the evaluation as well as compliance on the different interventions using tracker surveys 1 & 2.	November 2014
Report on Uptake and Maximizing participation for marginalized women	Evaluation of design calibration in SFM 2013-14 intervention. This report will provide details on interventions carried out as part of design calliberations to SFM 2013-14 scheme and give details on what impact each intervention had on take-up of SFM 2013-14 training.	December 2014
Report on Design of Market Linkage Intervention	Report on the motivation, objectives, evaluation design and design of a market linkage intervention which will be implemented with graduates of PSDF's SFM 2013-14 scheme. The document will also provide findings from baseline evidence on the evaluation sample as well as details on the sampling methodology for the evaluation.	January 2015
Report on Design of Job Matching Intervention	Strategy document on the motivation, objectives, evaluation design for evaluating a job placement intervention as design calliberations to PSDF's Skills for Jobs (SFJ) scheme. Using baseline evidence, this report will detail on the design of the intervention itself as well as the sampling methodology for the evaluation.	February 2015
Report on Design of BPSV Scheme	Design concept of the Big Push on Skills for Villages (BPSV) scheme using in-depth baseline evidence and analysis. The document will detail the motivation behind the intervention and evaluation, it's objectives, and evaluation methodology for evaluating the intervention.	May 2015

	This report will also provide details on the intervention design and the sampling methodogloy for conducting the evaluation.

Table 4 provides details of the proposed dissemination activities planned during the evaluation phase 1.

Table 4: Dissemination Activity for Evaluation Phase 1

Dissemination Activity	Detail	Expected Timelines
Event to disseminate results from the baseline reports	Event on which results from baseline reports and their implications for the design of skills programs will be disseminated. The baseline reports will also be released at this event. We hope to hold this activity in collaboration with PSDF and the Institute of Development and Economic Alternatives (IDEAS).	January 2015 2014
Dialogue on designing effective skills programs for women	This event will use the results from the *SFM* RCTs and the baseline surveys to discuss challenges and solutions that ensure equal opportunity to attend skills training for women. We hope to hold this event in collaboration with PSDF, the National Commission for the Status of Women and the Institute of Development and Economic Alternatives (IDEAS).	May 2016

Budget in Phase 1

The budget for Phase 1 is provided as a separate document. The budget document includes all costs of intervention and monitoring applicable to the program until the end of June 2015.

The contracting between DFID and CERP for Phase 1 of the Evaluation Period will be output based i.e. budget payments from DFID will be linked to proposed outputs to be delivered by CERP during Phase 1. Proposed outputs along with corresponding budget payments are provided as part of Appendix A.

Phase 2 (July 2015-March 2017) Work Plan

This sub-section relates to data collection activities and outputs to be delivered under the Phase 2 of the Evaluation Period i.e. from July 2015 to March 2017.

Interventions in Phase 2

Evaluation 2: *Market-linkage evaluation with rural women (large sample)*

As mentioned above, evaluation 2 will be conducted on a larger sample in phase 2 to ensure that we have ample power to detect impacts of the intervention on economic and non-economic outcomes of female trainees. In phase 2, market linkages will be provided for the graduates of a set of 32 random centers from *SFM* (2013-14) scheme. The expectation is that 10 trainees per centre (or 50% of the class) will avail this facility after successful completion of training.

We expect to roll out the second market linkage component from August 2015 to January 2016. The choice of which model to employ will be finalized based on findings from the first market linkage component and in collaboration with PSDF.

Tracker Surveys in Phase 2

Individual outcomes will be measured through tracker surveys related to the *SFM, SFJ* and *BPSV* evaluations. The purpose of these trackers is to provide evidence of how outcomes hold-up over time as well as enhance statistical precision in the quantitative analysis for end of programme evaluation. The description, proposed time and sample size for the tracker to be undertaken in Phase 2 is given below

- Tracker 6 will be conducted in January – March 2016 with a sample of 21,200 households and 1000 *SFJ* employers. This will track post-training outcomes for 13,700 individuals in the *SFM* treatment and control sample; and 7,500 individuals in the *SFJ* treatment and control samples.

- Round 7 of the tracker surveys will be conducted in May – August 2016 with a sample of 20,200 households and 1,000 *SFJ* employers. This will track post-training outcomes for 7,500 individuals in the *SFJ* treatment and control samples and 12,700 in the *BPSV* sample.

CERP will raise funding for the 1 round of tracker with 13,700 *SFM* treatment and control sample households in tracker 6; and for 1 round of tracker with 12,700 *BPSV* sample households. This equals PKR 32.35 million which is 160% of the budget that is being requested from DfID to support the tracker survey activity in Phase 2.

End line Surveys in Phase 2

End line surveys related to SFM, *BPSV* and *SFJ* will be conducted during Phase 2. The end line survey with the sample of 10,700 *SFM* CERP voucher holding households will be conducted in April-May 2016; whereas end line survey with the sample of 3,500 SFJ CERP voucher holding households, and 12,700 BPSV households will be conducted in November 2016 - January 2017.

The advantage of the end-line surveys is that they will allow us to measures impact on household poverty and vulnerability and changes in outcomes of interest at the level of the household as well as spillovers within the household. These surveys are important as they will allow us to see how the effect of training and future employment affects outcomes of others in the household as well.

Tracking outcomes at this frequency will allow us to get sufficient power to detect impact of program on economic and non-economic outcomes of interest at the level of the individual and household.

In the case of *BPSV* the end line survey promises the additional benefit of providing village-level outcomes of interest (such as GDP or aggregate income, savings, asset creation etc.) as the in-depth sample is a representative sample of these villages. This will allow us to see to what impact village-level human capital formation has on village-level outcomes and how spillovers occur within the village.

CERP will raise funding for conducting the end line survey in half of the *BPSV* evaluation sample, i.e. approximately PKR 18.16 million.

Table 5: Work Plan Proposed for Phase 2

Work Plan for Phase 2 (July 2015-March 2017)	2015						2016												2017			
	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4
INTERVENTIONS																						
a) Roll-Out of market linkage intervention 2																						
SURVEYS																						
a) Tracker Surveys																						
vi. Tracker Survey 6 (HH sample = 21200; Employer sample = 1000)																						
vii. Tracker Survey 7 (HH sample = 20200; Employer sample = 1000)																						
b) End-Line Household Survey																						
i. SFM end-line survey (HH sample = 21200)																						
ii. SFJ end-line survey (HH sample = 3500)																						
iii. BPSV end-line survey (HH sample = 12700)																						
Reports																						
a) Impact Evalaution Report of SFM																						
b) Impact Evaluation Report of SFJ																						
c) Final evaluation report of SFJ																						
d) Report on Evaluation of BPSV																						
e) Synthesis Report																						

Expected Outputs and Dissemination Activities in Phase 2

Findings from the evaluations in Phase 2 will be provided in the form of notes and reports that are listed in Table 6. Dissemination activities are listed in Table 7.

Table 6: Publications by CERP Proposed for Evaluation Phase 2

Publications	Detail	Expected Timelines
Report on impact of SFJ (and its sectoral variants) Intervention	Report on evaluation of SFJ Intervention which will give the impact of training and job placement intervention on economic and non-economic individual level outcomes in SFJ (and its sectoral variants).	March-April 2016
Report on impact of SFM 2013-14 Intervention: Long run impact of training for rural women	Report on evaluation of SFM 2013-14 which will give the impact of training and market linkages intervention on economic and non-economic individual and households level outcomes	August 2016
Final Report on impact of SFJ (and its sectoral variants) Intervention	Report on evaluation of SFJ Intervention which will give the impact of training and job placement intervention on economic and non-economic individual level and household level outcomes in SFJ (and its sectoral variants).	March 2017
Report on evaluation of BPSV Intervention	Report on the evaluation of BPSV Intervention which will give the impact of the intervention on outcomes at the household and village level.	March 2017
Synthesis Report that summarizes findings and recommendations of all RCT-based impact evaluations	Report will provide an analysis of relative cost-effectiveness and value of money associated with interventions evaluated during this phase. It will also provide an analysis of sustainability after one-year of the intervention. It will also report on one-year impact of *SFJ* (and its sectoral variants) and the job matching intervention by drawing on data from the 2nd *SFJ* tracker and the end line survey.	March 2017

Table 7 provides details of the proposed dissemination activities planned during the evaluation phase 2.

Table 7: Dissemination Activity Proposed for Evaluation Phase 2

Dissemination Activity	Detail	Expected Timelines
Event to disseminate the results of *SFM* plus market linkages intervention	This event will use the results from the evaluation of the *SFM* program and the market linkages intervention to discuss challenges in designing effective skills programs for women in high poverty districts and the potential of these programs. We hope to hold this event in collaboration with PSDF, the National Commission for the Status of Women and the Institute of Development and Economic Alternatives (IDEAS).	March 2016
Event to disseminate the results of *SFJ* plus job matching interventions	This event will use the results from the evaluation of the *SFJ* program and the job matching intervention to discuss challenges in designing effective skills programs for urban areas and the potential of these programs. We hope to hold this activity in collaboration with PSDF and the Institute of Development and Economic Alternatives (IDEAS).	October 2016
Event to disseminate the results of *BPSV*	This event will discuss challenges in designing effective skills programs for agriculture and livestock and the potential impact of such programs. We hope to hold this activity in collaboration with the PSDF, the *BPSV* consortium and the Institute of Development and Economic Alternatives (IDEAS).	March 2017

Budget in Phase 2

The budget for Phase 2 is provided as a separate document. The budget document includes all costs of proposed intervention and monitoring applicable to the program from July 2015 to March 2017.

5. The Project Team

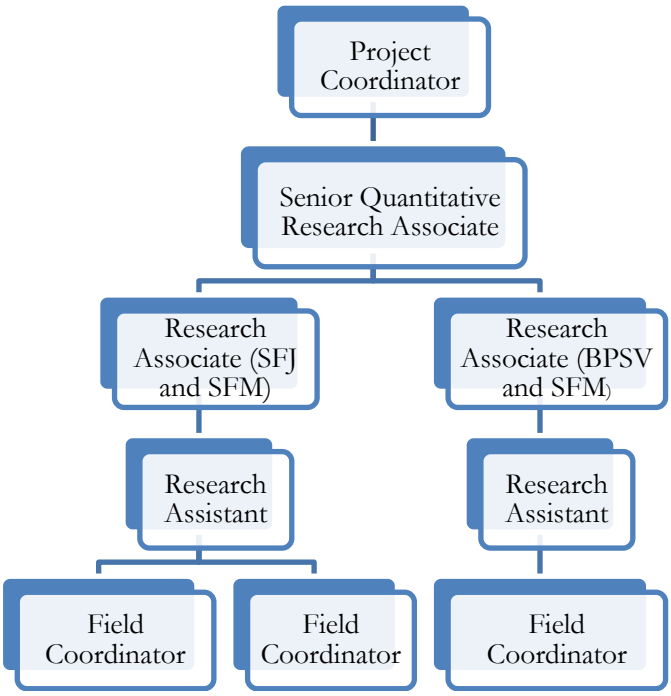
The project team will be headed by a Project Coordinator (with a post-graduate degree in management, economics or public policy with managerial experience) who will be supported by a Senior Quantitative Research Associate (with a post-graduate degree in economics). The current Project Coordinator has a post-graduate degree in Economics from the Lahore University of Management Sciences (LUMS). He has been associated with the project since its inception and has had approximately three years' experience running RCT-based impact evaluations and supporting baseline work. He has held the project coordinator position for two years and has developed an excellent working relationship with project partners. He has also demonstrated the ability to assure the quality of the RCTs and manage a large and diverse team of researchers. The Senior Associate is required because extremely rigorous quantitative work is required in the project and includes power calculations, sampling, running regression results, field monitoring, audit and analysis reports. This senior team will report to the CERP PIs.

The project PIs include two local PIs: Professors Ali Cheema (CERP, Institute of Development and Economic Alternatives and Lahore University of Management Sciences (LUMS)) and Farooq Naseer (CERP and LUMS); and two US-based PIs: Professors Asim Khwaja (CERP and Harvard) and Jacob Shapiro (CERP and Princeton). The PIs take all key decisions on the project jointly, though each has responsibility for particular aspects of the project. The local PIs provide oversight on the implementation of the project and are directly involved in the interface with all stakeholders. The US-based PIs become part of this engagement through weekly conference calls that are structured around a weekly call document that captures progress on a weekly basis, identifies areas for discussion and decisions and proposes solutions that are decided by all PIs jointly. The local PIs provide oversight on compliance with the interventions including monitoring of implementation; ensure that agreements are obtained with PSDF Board about evaluation design; monitor the progress of survey and qualitative field activity and are part of the field activity itself. They are also responsible for providing oversight to the general project administration and management. Local PIs attend all strategic level meetings with PSDF and other stakeholders. The US-based PIs take the lead in providing input into the design of the evaluation methodology and are jointly involved in statistical analysis for sampling and results and the writing of results and reports. They are also an important conduit of presenting this work internationally to academic peers and international policymakers, which gives depth to the peer review process. US-based PIs have also been deeply involved in the development of field instruments including survey questionnaires and revisions to these instruments after taking into account the results from the pilot exercise. The US-based PIs spend approximately one month in a year in Lahore during which time they have structured engagements with PSDF and its board members. They have also virtually participated in strategic meetings with PSDF through video conferencing facilities. Additionally, Feyza Bhatti of the University of Edinburgh and Rabea Malik of the Institute of Development and Economic Alternatives will provide targeted technical gender expertise and oversight of the qualitative field instruments and methodologies.

Two Research Associates (having a post-graduate degree in economics) will work with this senior team and they will support the Senior Quantitative Associate on quantitative tasks and the coordinator on field related tasks. They will collaborate on the *SFM* (Evaluation 1) and the market linkage (Evaluation 2) evaluations but will be individually responsible for *SFJ* (Evaluation 3) and *BPSV* (Evaluation 4) evaluations, respectively. Each Associate will be supported by a research

assistant (with a graduate degree in a relevant social science). The field is managed by 3 field coordinators two of whom will be placed in the pilot districts and one will be responsible for the expansions districts.

Figure 1: Organogram of the Project Team



6. Impact Summary

The impact of the Evaluation Phase of the project is expected to be multi-fold. Rigorous evaluative evidence will provide tangible results of the impact of PSDF training on log frame outputs and outcomes. We will not only measure the impact on increased income and consumption but also on non-economic returns to training. These include state engagement, civic participation, empowerment and mental health, empowerment, and externalities (positive and/or negative) on other household members, especially women and children. Our comprehensive panel surveys and qualitative focus groups capture these impacts. Impact will be measured for males and females and for rural and urban citizens. These results will inform the PSDF Board and management and provide essential evidence for the design of their future strategy. The results will provide the necessary evidence that will enable the PSDF Board, Government of Punjab and DfID to assess the cost-effectiveness and value for money associated with different interventions. This will allow them to make an evidence-based decision the composition of the portfolio of interventions that should continue to be supported through public investments.

Rigorous evidence-based research on design challenges will identify first-order challenges that are constraining access and muting the returns to skills training in a LIC context. Impact evaluations of calibrated interventions will provide rigorous evidence on the net returns associated with different design solutions to these challenges. This will directly impact development thinking globally about cost effective solutions that raise access as well as the returns to training in a LIC context. In particular, the evaluation phase will provide evidence about the cost effectiveness of a range of interventions (village-based training, safe and reliable transportation, information provision, financial and credit constraints and community mobilization) designed to mitigate access problems for women in a social context of low mobility. It will also provide evidence on the increase in net returns associated with pairing in-class training with market linkage interventions for women in this context.

Evidence will also be provided on the additional returns that accrue if in-class training is paired with job matching interventions. The evaluation phase will also provide evidence on whether saturating villages with "big push" training in agriculture, livestock and related skills can convert villages into engines of growth. In this case, we also have opportunity to rigorously measure the impact of skills training on village level GDP and productivity, gains to households at different points in the income distribution, and the extent of spillovers that result from interventions.

This unique collaboration between policymakers, providers and researchers should serve as a model for mainstreaming evidence-based policymaking. It would be important to disseminate the model of "smart policy design" that has evolved out of this collaboration, which cost-effectively evaluates interventions designed to address critical program challenges in their early design stages and offers the opportunity for recalibration that promises higher returns.

7. Wider Impact and Sustainability

Dissemination Strategy

One of the principal strengths of this project is that it is uniquely situated to allow for quick and direct uptake from research findings to policy. Our strategy is based on the premise that research uptake happens most effectively when the appropriate policy stakeholders and key audiences are directly engaged in the entire process; this is precisely what we have in place already with the Punjab Skills Development Fund (PSDF), and the Government of Punjab. PSDF represents an unusual partnership between the government (bureaucracy and political establishment), donors (i.e. DfID-Pakistan), civil society, the private sector (the board has membership and guidance from both non-governmental and private sector organizations active in this space) and researchers that is highly relevant in the context. Their involvement at all stages of the project–from gathering of evidence to inform design to assessing the short- and long-term socio-economic impacts of interventions–allows them to experience the full process through which evidence is used to inform policy. This generates not only specific knowledge about the program but also helps build a broader appreciation of the evidence based decision making process.

As evidence of this project's potential to influence policy we note that the results of our past work are already informing PSDF policy and our reports have garnered substantial attention. Our household and employer reports based on representative samples of close to 11,000 households and 7,500 employers helped inform program design in a variety of ways, including: better matching course offerings and content to trainee and employer demand; adding basic literacy and numeracy modules to the courses so as not to exclude poorer and more marginalized individuals; and introducing job matching & placement services. Highlighting the importance of distance as an access issue led to a pilot intervention that demonstrated the value of village-based training and has greatly informed the current set of interventions outlined in this proposal. Throughout this proven example of cross-sectoral collaboration, PSDF has not only shown an understanding of how to use research in policy-making, but their new initiatives have been directly and explicitly informed by our evidence. They have demonstrated the value of the research in new program designs and have cultivated an organizational culture of evidence-based learning.

Additionally, PSDF engages multiple public, non-governmental and private training service providers (TSPs) to deliver capacity building activities rather than doing them in house. As a result, as researchers, we also have started building strong relationships with these providers and our interventions will be done in close conjunction with them. This research funding will help PSDF and TSPs make critical decisions on how to allocate scarce funds during the scale-up, as well as establish precedent and inform decision-making for other actors (both governmental and non-governmental) that are seeking to reach rural women and are constrained by distance.

We have also established a cooperative agreement with the National Commission on the Status of Women (NCSW)–the premier agency responsible for examination of policies for women's development and gender equality–which will serve as an additional platform through which to disseminate findings to a broader policy making audience. We will seek similar agreements with premier provincial government departments, non-government organizations and private sector organizations engaged in sectors that will be the focus of evaluations.

Dissemination will happen through local and international seminars and conferences and Policy Dialogues. Policy Dialogues will be organized to enable stakeholders to grapple with the evidence and work together to develop effective policy solutions through interactive problem solving and strategic coordination. In addition, we will produce policy reports and more accessible policy briefs that will also disseminate and document the results. A special effort will be made to share results with relevant policymakers and stakeholders from other provinces. Details of these dissemination activities are given in section 7.

Because of our and PSDF's sustained relationships with these stakeholders, we can be confident that policy makers and implementers will be receptive to the research output we produce. The questions that we ask have, in part, been directly influenced by their experiences and the types of questions they are asking. Because we have demonstrated through prior engagement that research findings do improve policymaking, we can leverage that precedent to encourage future uptake and receptiveness of the latest round of results. Keeping in mind that broader public buy-in is also crucial to research uptake, existing connections with traditional and social media sources will ensure public dissemination of results. CERP has recently begun implementation of the DfID-funded Building Capacity to Use Research Evidence (BCURE) program in Pakistan (in addition to Afghanistan and India) in collaboration with Evidence for Policy Design (EPOD), an initiative at Harvard Kennedy School. EPOD (http://epod.cid.harvard.edu/) has a strong track record of utilizing rigorous research to inform public policy, as well as regularly engaging policy makers and strengthening their ability to utilize evidence in policy-making. Collaboration with EPOD will also ensure global outreach of the findings of the evaluation phase.

Capacity Building

The fact that evaluation design, survey design, data collection and data analysis is done at the Center for Economic Research Pakistan (CERP) means that this is very much a program being spearheaded by a local organization that has built capacity for international standard evidence-based policy research. Pakistani PIs are very much at the forefront of this research, with US-based PIs providing design input, intellectual support, and capacity building opportunities as needed. This unique collaboration is building institutional capacity to undertake rigorous impact evaluations in Pakistan and creating opportunities for emerging Pakistani leaders and scholars to form enduring ties to the international development community. Apart from the PIs CERP researchers and staff are deeply involved in all stages of the evaluation, from intervention design, to sampling, to intervention monitoring, to statistical analysis, to drafting reports and academic papers.

The CERP field team consists of 8 research staff (section 9), a steady flow of local interns, and additional field staff and surveyors, all whom gain valuable experience designing and implementing research in collaboration with policy makers. These regular interactions represent the beginning of mutually beneficial research policy relationships that will last well beyond the timeline of this program. Project researchers have already won Fulbright awards (4 research assistants and associates) and PhD placements (2 project researchers) in top international graduate programs in economics. We very much hope that the experience of these researchers with the project will result in a number of high quality academic research papers on labor market, human capital and gender, human capital issues in Pakistan.

Policy-makers in Punjab—both in PSDF, Industries and Planning and Development— as well as the Chairperson of the National Commission on the Status of Women (NCSW) have been involved in the program from the beginning and will continue to play a key role in its design and execution. Their consistent involvement is critical the success of the program and to facilitate uptake of the results as described above. PSDF will continue to build their own capacity to use evidence in policy design as they lead local efforts to move from anecdote- to data-driven decision making.

The research team conducts regular meetings with PSDF Board members, who represent a broad cross-section of policy stakeholders in Pakistan, provides a rich view of their needs and perceptions. Feedback from these meetings informs program design in both directions: the policy makers are learning how to design for evaluation and the researchers are learning about the various constraints policy makers must balance. These interactions have led to numerous minor improvements in both program and evaluation design. As discussed above, this partnership between the locally-driven research team and PSDF represents a unique collaboration that ensures research uptake but also places Pakistani policy makers and researchers at the forefront of the research policy engagement.

The training service providers and other private, NGO and civil society partners will be directly involved in the implementation of meaningful trainings, skills that they will be able to replicate even beyond the scope of this program. Building their capacity to use research to design programs and interventions will also be valuable since presumably not all programs they implement will be through official government funding mechanisms. Additionally, even donor (DfID) local staff will have exposure to the implementation and research sides of the program, gaining valuable skills as they see the evaluation being conducted and monitor its rollout.

Finally, by virtue of the program design, local residents in 14 districts in Punjab where access, poverty and productivity issues are salient will have an opportunity to gain valuable and tangible skills. We have also seen that resident preferences and views are utilized by policy makers as part of the decision-making process because they understand the importance of getting such info from the public in the first place.

8. Building on the Baseline Phase

The Evaluation Phase outlined in Section 2 of this proposal builds on the Baseline Phase of the project which has already contributed substantively to informing and recalibrating the design of the PSDF program. It has identified and disseminated design-relevant challenges related to skills training and livestock and dairy development through rigorous context-specific evidence on the supply and demand sides of the skills, labor, livestock and dairy markets. It has also provided a comprehensive mapping of the existing skills landscape in the pilot districts that has informed the design of the menu of trades offered by PSDF. It has produced a baseline for log frame monitoring and for the proposed impact evaluations of PEOP interventions.

This rich context-specific evidence is the result of large-scale survey data on households, villages, and employers that has been collected as part of the baseline phase in the PEOP pilot districts. Specifically, CERP oversaw household survey activities among a non-in-depth sample of 10,495 households in 429 villages and 280 urban clusters, and among an in-depth sample of 21,000 households in 149 villages. A village census was also conducted in 149 rural population sampling units (PSUs), and an employers' survey covered a district representative sample of approximately 6,200 employers. Livestock supply side surveys were also undertaken, as well as a survey of PSDF training course graduates to assess post-training employability.

In addition, CERP's engagement with the PSDF Board during the baseline phase has evolved a model of smart policy design that is built around continuous cost-effective learning through small-scale phased evaluations of design calibrations and evidence-based calibration of the program. This model has demonstrated early success in the baseline phase and has tremendous potential to effectively calibrate the PSDF program during the evaluation phase. This model can be easily replicated in other public sector and DFID programs and has the potential to create tremendous value.

Three pilot evaluations relating to the main PSDF programs have been conducted as part of the baseline phase. A pilot evaluation of the SFE scheme was conducted to understand the variations in training uptake within the target population and to provide evidence on the correlates of low uptake. Pilot evaluations were also used to calibrate early stage PSDF interventions designed to increase uptake among rural women and male income earners in the target population. PSDF experimented with a model of village-based training as part of the SFM scheme to mitigate distance related constraints for women and the expectation was that easing this constraint would have a large impact on uptake among women. As part of the SFJ scheme, a pilot evaluation was designed under which PSDF increased stipends to offset the risk of training for male income earners that resulted from having to give up jobs to take up training. The expectation was that this intervention has the potential to increase uptake among this group as it compensated for this risk. Finally, the baseline phase also emphasized that agri-livestock skills are in high demand but cost-effective provision may require a more "big push" style delivery that leverages the experience of successful players. The baseline phase revealed the mismatch between the demand for agriculture and livestock skills and the inelastic supply response. The recognition of this mismatch has prompted PSDF to involve a consortium of private-sector progressive farmers and companies which have adopted frontier practices to help seed an effective supply response for these types of training.

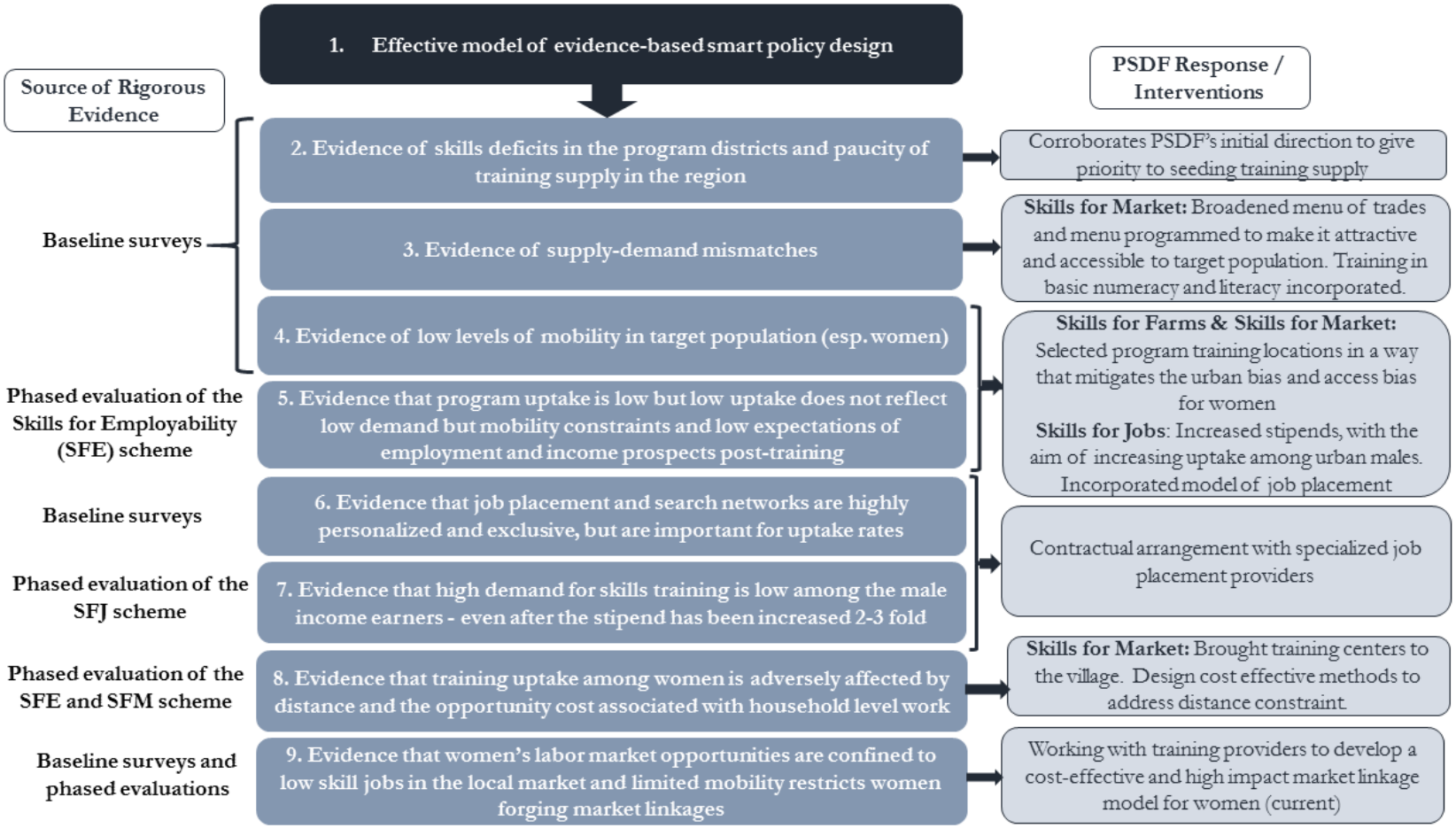
Evidence from the baseline phase revealed that increasing the impact of PSDF programs in the PEOP districts requires addressing two interrelated issues; increasing uptake of training

opportunities among the target population and increasing the "value-add" of training courses. Both of these factors affect measured impact, which is a product of the uptake rate and training value added. Both of these inputs are, in turn, constrained by a series of "first order" design challenges. Rich evidence on factors constraining uptake and the value-added of training has been produced by CERP during the baseline AGA phase. The evidence produced during the baseline phase (for details see Appendix D) has had direct program impact in a number of areas, which are summarized in Figure 2 below.[33]

The Evaluation Phase will build on the foundation established in the Baseline Phase, with an ultimate goal of rigorously evaluating the socio-economic impact on beneficiary households and communities through a series of interventions that have been redesigned in light of the evidence gained and lessons learned from the baseline phase.

[33] Figure 2 summarizes the key findings from the baseline phase, gives the source of the evidence and provide details about how this evidence was used by PSDF to inform and calibrate design.

Figure 2: Summary of Impact



9. Ethical Considerations

The research and evaluations conducted during the evaluation phase will uphold the following principles, which have been pursued during the baseline phase as well:

1. **Ethics Approvals will be obtained from institutions with reputed and well-developed ethics approval systems:** All work conducted by CERP on the PEOP project has already taken place with the approval of either Harvard or Princeton's Institutional Review Boards, widely accepted to be amongst the most comprehensive human subjects review boards in the world. We will continue to submit all necessary protocols to the Harvard IRB to ensure human subjects compliance.

2. **Research and evaluation is relevant and high quality with clear developmental and practical value:** CERP has taken a number of steps to ensure relevance, quality and practical value.
    a. The **relevance** of the evaluations is ensured because the interventions being evaluated form the core of a large-scale skills development program that is being supported by public funds. Furthermore, the innovative interventions are well-grounded in context-specific evidence that was produced during the baseline phase. The evaluations are being undertaken to empirically validate the theory of change that underlies the PEOP skills program as there is a clear gap in knowledge in this regard. They are also being undertaken to understand the additional impact associated with addressing different context-specific constraints that are adversely impacting access and economic and non-economic outcomes. There is a clear gap in knowledge in this regard. For example, there are few studies that have rigorously evaluated access costs and impact of skills training on the general population and its different segments as most studies have focused on the selected population that show up for training.
    b. The **quality** of the evaluations is ensured through the following processes. First, evaluations designs are presented to the PSDF Board, shared with DfID and presented at reputed academic conferences to get peer feedback. In addition, for evaluations 1-3 we have successfully won supplementary funds through reputed competitive research grants[34] with a peer review process. We have voluntarily competed for funding that entails a rigorous quality review and selection process precisely because we are committed to quality reviews as the CERP PIs are incentivized to further the body of scientific knowledge—not to promote a particular program, policy or organization—as one of their main objectives is to contribute to the academic literature in the area of development by publishing in reputable peer reviewed academic journals.
    c. The **practical value** of the evaluations is ensured as it is dealing with core interventions of a major public policy program in a critical developing economy.

[34] These grants are: Growth and Labour Markets in Low Income Countries Programme (GLM LIC); IZA-DFID; and Growth and Economic Opportunities for Women (GrOW); IDRC, DFID and The Hewlett Foundation

3. **Harm is avoided to participants in the studies and confidentiality of information, privacy and anonymity of study participants is insured:**
    a. Strict protocols are put in place that safeguard the privacy and confidentiality of respondents, all personally identifiable information is only saved in encrypted password-protected files. Individual respondent's answers to survey questions are not shared with PSDF or training service providers under any conditions. CERP organizes a helpline during the implementation of its evaluation protocols and any complaints about program services received through the helpline are communicated only with PSDF with the consent of the respondent, under strict confidentiality. Respondents raising issues and complaints during the survey process are encouraged to call the helpline and record their complaint or SMS the helpline with a request to call the respondent back.
    b. As with any program with limited available resources, benefits and/or slots, the selection of those who will be treated presents ethical questions. By design, the overall program provides access to training classes and other benefits to some pre-target populations but is unable to do so for every member of those populations due to limited resources. Our role in this process has been to ensure that populations that may be marginalized on account of their socio-economic status are given careful attention through program design. Moreover, to the extent that limited slots and benefits have to be allocated across people, the traditional method is to allow the training service providers to simply choose whomever they want to enroll with little or no checks on who they are selecting and who is being excluded. We have successfully argued against this process during the SFM evaluation and have been successful in getting the training providers to record background information on ALL applicants, and to the extent that one is unable to differentiate between their needs, use more equitable methods like public lotteries to allocate the slots and benefits, thereby ensuring transparency and equal chance to all motivated parties. We believe this system to be significantly fairer than the usual local systems of resource and service allocation that can often be based on local relationships or subjective needs assessments by biased local leaders. Furthermore, we have requested PSDF to offer multiple rounds so that those who are unsuccessful in the initial ballot can be accommodated in later rounds of skills training courses.
    c. All benefits and services that are made available as part of an evaluation are presented to the prospective trainess transparently, both in writing and verbally. The receipt of this information is recorded as part of the intervention roll out protocol and the data is back-checked and verified on a sample basis. Furthermore, consent is obtained that the participant understands the benefits being offered to them and others and is willing to participate in the evaluation. Verbal communication is an integral part of the protocol as we are dealing with a population in which literacy is low. An independent call centre is set up that calls beneficiaries to check if the benefits are being received on time and as per the terms of the program and this information is triangulated at the village level through random field checks. These protocols are part of the treatment compliance activity.
    d. Regarding the actual interventions proposed, preliminary surveys show that skill building services are highly sought after and present almost no risk to those participating. Nevertheless each participant is provided detailed information about the training courses including; not only the range of course options, locations and dates available, but also what the likely skills gained would be.

e. Members of the control group are not excluded from attending training and can in principle access the standard mechanisms of information dissemination about the training. Members of the control group in the general population are, however, not provide information as intensely as the treatment group. Again we use the mechanism of a public ballot that randomizes on villages and not on specific populations in villages to create the control and treatment groups in the general population (voucher-based) RCTs. In the case of the oversubscription design, we are recommending a design with multiple rounds so that members of the control group can access training at a later date and are committed to randomizing using a public ballot.

4. **Participation in the research and evaluation is voluntary and free from external pressure:** We follow strict protocols to provide all information about the treatments being offered in writing and verbally. During the intervention roll-out, a consent script is read and a household is only enrolled in the evaluation if consent is given. Verbal consent will avoid complications that arise from working with less literate populations. This activity is conducted with trainees within their household and away from community pressure.

5. **Protocols are designed to respect cultural sensitivities:** This is ensured by piloting these protocols and obtaining feedback on them through qualitative fieldwork. This is an important part of the treatment compliance activity.

6. **CERP is committed to publication and communication of all evaluations and research studies:** Full methodological details are made public through policy reports and as part of the academic publications that will results from the project. The latter commitment ensures that the methodology is peer-reviewed to meet the highest standards of quality. We are committed to eventually make anonymized data and code available for any external party to be able to replicate the results. We will also publish the intended analysis on the American Economic Associate registry as an additional measure of commitment to peer review and transparency.

7. **All evaluations have particular emphasis on ensuring participation of women and socially excluded groups:** The interventions we are evaluating are designed to improve access among women and socially excluded groups and compliance with this objective will be closely monitored as part of the treatment compliance activity and reported transparently.

8. **The evaluation will be independent of the program under study:** We detail reasons why there should be confidence that the evaluation will be completely independent. However, we support the recommendation of the reviewer to establish an Evaluation Steering Committee to strengthen the institutional basis for ensuring that the evaluation remains truly external and independent and to ensure transparency of process.
    a. The CERP PIs do not gain any monetary benefit from this project or with PSDF achieving a particular outcome. The decision of the CERP PI's to fully donate the cost of their time is in order to dispel any suggestions about lack of independence. Furthermore, less than 6% of the overall budget will accrue to CERP in the form of indirect costs, which are a compensation for overheads and management time of CERP administration. This amount is significantly less than the contribution of the CERP PI time that is being donated. It needs to be appreciated that even these

indirect costs were donated by CERP during the baseline phase in order to ensure complete financial independence. In addition, as stated earlier we have subjected ourselves to the rigors of internationally reputable and competitive research grants that have a strong peer review process associated with them. Furthermore, being academics at reputable universities and research institutions we are incentivized to contribute to the creation of scientific knowledge by providing precise estimates of the returns to different interventions and publishing the evaluation design, methodology and results in reputable peer reviewed journals that require data and codes to be made public to external reviewers for the purposes of replication of results.

b. The engagement process set up during the baseline phase was structured to ensure academic autonomy for the research produced by CERP and an arms-length relationship has been maintained between PSDF and CERP during the production of research. The research agenda was informed by theory and review of literature and incorporated suggestions from PSDF (including feedback from their stakeholders), DfID and Government of Punjab and this agenda was transparently documented in the form of the baseline phase proposal which was accepted at three levels: (a) the PSDF Board; (b) DfID after undergoing a rigorous review process and (c) a higher PEOP Steering Committee with wider representation from Government of Punjab, DfID, private sector and academia. CERP designed survey instruments and methodologies to undertake this work and hired a survey company through a transparent and public procurement process (with independent members being part of the selection committee) to undertake the survey. The initial engagement with PSDF was essential to ensure the policy relevance of the baseline phase and to ensure context relevance. The results were produced by CERP in the form of notes and reports that were shared with DfID, the PSDF Board and the PEOP Steering Committee. In addition, the results have been publicly presented in conferences and seminars organized by the Lahore University of Management Sciences; the International Growth Centre; Harvard; Princeton; DfID and by IZA-DfID's Growth and Labor Markets in Low Income Countries Program. This evidence was presented to the PSDF as an input into their intervention design exercise as well. This was built as a necessary part of the baseline phase engagement in order to ensure that the results directly impacted design thinking and there was ownership of the results by the PSDF Board. During this period, PSDF has commissioned independent research and surveys and CERP has not participated in these bids in order to avoid concerns about conflict of interest. The evaluation phase will continue this arms-length working relationship between CERP and the PSDF Board.

c. CERP does not have involvement in the design of the core PSDF training schemes, such as *SFM*, *SFJ*, *Skills for Garments, Skills for Farms, Skills for Employability* and *Skills for Biogas*. These schemes have been designed independently by PSDF and their technical consultants with approval of the PSDF Board. CERP's role has been to provide evidence about margins that the current design of these schemes are ignoring and the impact this has had on realized demand, selection of different social, income and gender groups and potentially on returns. One of the main purposes of the evaluation phase is to evaluate the impact of these schemes.

d. However, CERP *is* providing evidence-based feedback into the design of innovative interventions that are complementary to training such as market linkages and job placement. The design of these interventions is being proposed by providers with

evidence-based feedback from CERP's trainee, household, community and market surveys. The timely availability of evidence ensures that rich analysis is available for design, which promises high benefits in terms of relevance and basing designing on an analysis of context-specific constraints and opportunities. It needs to be clarified that CERP is not advocating particular development approaches while providing technical assistance and is not wedded to specific positions. Its main responsibility is to provide timely evidence using primary data on the dimensions of impact that promise returns and design specifications that are unlikely to work with target beneficiaries. One complication is that CERP has been asked by PSDF to manage the access interventions for *SFM* and the *Market Linkage* interventions because financing these interventions through PSDF imposes significant overheads as they will have to set-up elaborate open bidding processes which will be highly costly relative to the scale of the interventions. It is for these reasons that CERP has also been asked to provide oversight into the development and use of the online application registration system that is needed for the *SFJ* oversubscription design. CERP has assumed this role reluctantly with consent from DfID and would ideally like all these interventions to be funded through the PSDF budget.

e. The ideal of division of responsibility in CERP's view is that PSDF assumes responsibility for the design of interventions (with evidence-based design feedback from CERP) and the procurement of providers that implement these interventions. CERP's main responsibilities should be to ensure that the design is structured in ways that make rigorous evaluation possible and to conduct the evaluations. This will entail ensuring that there are sufficient treatment and control units to implement an RCT-based methodology; treatments are sufficiently well-specified to be reproduced if they work; and implementation is based on a well-specified and documented theory of change such that the evaluation contributes to general knowledge. Ultimately CERP's main role is that each of these interventions is rigorously evaluated in a manner that is transparent and replicable. Finding ways to achieve the ideal division of responsibilities and ensuring that the design is structured to make rigorous evaluation possible are responsibilities that the proposed Steering Committee can assume.

f. CERP welcomes the proposal of setting up a Steering Committee that ensures the independence of the evaluation process and agenda. In addition, we feel strongly that CERP's commitment to ensuring verifiability of evaluation design, methodology and results through publishing results in reputable academic peer reviewed journals and the intended analysis on the American Economic Association registry will ensure independence of the evaluation.

10. Risk and Risk Mitigation

Given the nature of the researcher and policy maker collaboration we believe that any political, reputational and operational risks are minimal. Our team is almost entirely made up of local researchers and practitioners who are well-integrated into local society. The existing strong relationship with PSDF, NCSW and the Punjab Planning and Development Board—all pivotal agencies on skills and planning in Punjab—also legitimizes our presence in communities and in the broader policy sphere within which we will be operating. Adding to the feasibility of the project as a whole and risk mitigation is the fact that the Board of Directors of PSDF has a broad array of civil society, political, public and private members, each of which ensures that all stakeholder views are represented and considered. Furthermore, the content of the evaluations presented in this proposal have been approved by the PSDF Board.

There is minimum risk associated with the *SFM* (2013-14) evaluation as the design has been finalized and approved by the Board and it has been successfully contracted in the field with an enthusiastic response from the community and potential trainees. The fact that PSDF and CERP have been able to successfully contract the services required for the implementation of a RCT-based impact evaluation at this scale even though it involved 6 training providers in 3 different districts suggests that sufficient capacity has been developed at CERP to roll out large-scale impact evaluations. A robust monitoring and audit plan is being put in place to ensure post-contractual compliance with the roll out of the evaluation.

There is minimum risk associated with the evaluation of the *SFJ* scheme. The methodology being used has been successfully used in RCT-based vocational training impact evaluations in Latin America and Africa. The proposed oversubscription methodology has been successfully piloted as part of the *SFM* (2013-14) evaluation as well. Therefore we have the learning and the experience to undertake an evaluation of the *SFJ* scheme using the oversubscription design.

There is low to medium risk associated with the efficacy of the design phase of the *SFM* market linkage intervention and the *SFJ* job matching intervention. In the case of market linkages we have received developed proposals from five potential providers and are in the process of evaluating these proposals. One of the five providers has experience with successful implementation of market linkages in the Pakistani context. Another provider has well-developed networks with branded retail outlets that are interested in carrying a dedicated line of products that are produced in high poverty districts.

The *SFJ* job matching intervention is building on efforts to register employers that have begun at PSDF and are showing positive signs. In order to mitigate this risk the design of *SFJ* job matching is being spearheaded by private sectors members of the PSDF Board and PSDF management who will use their networks to stimulate the creation of a registered database of potential employers. We anticipate working with training providers to help create a more refined report cards for trainees and given our past experience, do not anticipate this being a significant hurdle. We therefore also anticipate low risk in this regard.

BPSV has a somewhat higher risk than the other designs but has the promise of extremely high returns. The main risks are associated with developing a comprehensive readily-implementable training curriculum to provide frontier practices and skills and the seeding of cost-effective training providers to deliver the training. There is a risk that the proposed consortium of private-sector

companies and progressive farmers may not come together thereby making it difficult to design a comprehensive menu of frontier skills and practices. In order to mitigate these risks the design of *BPSV* is being spearheaded by private sectors members of the PSDF Board that have experience and depth of networks in these sectors. The depth of experience in the PSDF Board in these areas combined with their ownership of these design calibrations and CERP capacity to provide evidence-based technical assistance suggests that the risk is manageable. The PSDF procurement model has been extremely effective in seeding training suppliers and we are confident that it will provide successful results in this case as well.

There is minimum risk associated with the response rate of tracker surveys. Three trackers have been conducted up till now in which the average response rates have been approximately 92% and 89% in rural and urban samples respectively.

A potential risk in all these evaluations is implementation compliance by training service providers (TSPs). Our design protocol is complicated and requires the training service providers to check their usual processes in a number of ways. PSDF itself has very careful operational and financial oversight of TSPs; we will further mitigate these risks by instituting a robust real time monitoring and audit plan with our own field staff monitors progress at multiple levels, maintaining close contact with TSPs throughout the design and implementation phases and through regular communication with PSDF.

There are a series of potential risks associated with the use of the RCT methodology that we believe pose minimum risk on the whole. The biggest risk relates to external validity of the proposed evaluations. It should be recognized that this is a generic critique of quantitative and qualitative empirical studies and is not a risk that is specific to the RCT approach. However, there are a number of factors that lower this risk in our context. Firstly, the evaluation samples are representative of the general population and the selected population enrolling in training in the pilot districts. This sampling strategy ensures that our results are valid for the environment of the pilot districts. These districts represent a very large and meaningful population and the environment in these districts is similar to high poverty districts in other parts of Southern and Western Punjab and Sindh. Furthermore, the extension of the evaluations to the expansion districts allows us to draw additional representative samples of the selected population enrolling in training in the more developed districts and our results will be valid for the selected population in this environment. Taken together these results will provide valid insights for districts with similar populations in the province; in other provinces and in other South Asian countries.

Another risk typically associated with RCTs is that they may end up looking at limited outcomes and might have a difficult time uncovering all the causal channels. This risk is low in our context because evaluations are being designed to identify the different channels more precisely. We have also integrated structured qualitative fieldwork to enable us to obtain a better understanding of the causal channels underlying impact as well as identify the role of important non-intervention factors. Furthermore, we are collating comprehensive data on a whole range of outcomes that will enable us to look at the impact on a broad set of outcomes that go beyond the primary indictors specified in the log frame. This will allow us to better triangulate the existence of impact for different sets of outcomes.

There is also the risk that the results 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. This ownership has been demonstrated by PSDF during the baseline phase. In order to further address this risk we are pursuing a targeted dissemination approach with pivotal policy players such as the National Commission on the Status of Women, the Punjab Government as well as the Government of KP.

There is medium risk associated with the sustainable deployment of interventions such as market linkages and *BPSV* that are outside the range of current practices. One might therefore worry how sustainable these interventions will be in the absence of CERP's evaluation. This risk is mitigated by the fact that the PSDF Board has shown a commitment to scale-up these interventions if they are found to be successful and because the policy returns associated with these interventions are recognized to be extremely high by provincial governments and by non-government development organizations.

11. Research Team

The group of Principal Investigators from CERP engaging with PSDF for the PEOP skills interventions are:

- Dr. Ali Cheema (Lahore University of Management Sciences)
- Dr. Asim Ijaz Khwaja (Kennedy School, Harvard University)
- Dr. Farooq Naseer (Lahore University of Management Sciences)
- Dr. Jacob Shapiro (Woodrow Wilson School, Princeton University)

The above researchers bring significant experience working on a wide range of empirical projects. Their work is also featured in leading academic journals and popular media outlets. These projects include LEAPS (Learning and Educational Achievements of Punjab Schools) which in ongoing, multi-year survey to judge the educational outcomes of 2,000 households in 112 villages of the province of Punjab (Dr. Asim Khwaja); an evaluation of RSPN's Community Led Total Sanitation Project which was conducted for 8000 households in Punjab, Sindh and Gilgit-Baltistan (Dr. Farooq Naseer); an impact evaluation of introducing child-friendly schooling in 70 public sector schools in Islamabad which involved testing 2,000 children (Dr. Farooq Naseer); a village mapping and a full dataset of 1,000 schools (including testing scores of 7,300 students) and 28,000 households (across 126 villages) in the earthquake-affected areas of NWFP, Pakistan (Dr. Ali Cheema); a 1560 households survey in 35 villages of the district of Sargodha in Punjab to evaluate intergenerational structural inequalities between social groups (Dr. Ali Cheema & Dr. Farooq Naseer with funding support from Institute of Development Studies, Sussex University); 16,000 and 6,000 household nationally representative surveys in Pakistan to analyze the relationships between labour market outcomes and political behavior as well as a 3,000 person survey in the U.S. to validate various methods of using surveys to measure sensitive attitudes and behaviors (Dr. Jacob Shapiro).

Individual profiles of each of the above mentioned researchers are given below:

Dr. Ali Cheema is Associate Professor of Economics at the Lahore University of Management Sciences (LUMS), Pakistan, Research Fellow at the Institute of Development and Economic Alternatives (IDEAS) and one of the founding Directors of the Centre of Economic Research, Pakistan (CERP). His areas of interest include household and regional poverty, rural development, social protection and labor mobility. His research combines extensive field work with rigorous empirical analysis of community structures and household behavior. He has been extensively engaged with policy design and analysis with a particular focus on the high poverty districts of the Punjab province. Cheema has served as a member (2008-10) of the Chief Minister Punjab's Economic Advisory Council; a member (2014) of the Punjab Government's Growth Strategy Working Group; and a member (2008-10) of the Panel of Economists of the Planning Commission. He is a founding member of Stockholm Challenge Award winning portal the Relief Information System for Earthquakes Pakistan (RISEPak). He is a Rhodes scholar and received a BA in politics, philosophy, and economics from the University of Oxford and a PhD in Economics from the University of Cambridge.

Dr. Asim I. Khwaja is the Sumitomo-Foundation for Advanced Studies on International Development Professor of International Finance and Development at the Harvard Kennedy School of Government and co-Director of the Evidence for Policy Design (EPoD) research programme at

Harvard. He serves on the board of directors of the Poverty Action Lab (MIT), one of the leading policy evaluation organizations worldwide. His areas of interest include economic development, finance, education, political economy, institutions, and contract theory/mechanism design, focusing on Pakistan. His research combines extensive field work, rigorous empirical analysis and micro-economic theory to answer questions that are motivated by and engage with policy. Khwaja's research has been published in numerous peer-reviewed economics journals, including the *American Economic Review*, the *Journal of Development Economics* and the *Quarterly Journal of Economics*. Khwaja received BS degrees in economics and in mathematics with computer science from MIT and a PhD in economics from Harvard.

Dr. Farooq Naseer is Assistant Professor Economics at the Lahore University of Management Sciences (LUMS), Pakistan. He has previously worked on the impact evaluation of a pilot programme introducing child-friendly teaching methods in public sector schools. Over the past two years, he has worked on the issues of social stratification, poverty and inter-generational mobility using empirical data from detailed household surveys conducted in Sargodha district. The findings from this work have been used to inform various public policy debates and will be formally submitted as a report to the Planning and Development Department, Government of Punjab. He is a member of the LUMS Development Policy Research Center's (DPRC) Steering Committee. Naseer holds a BS in economics from LUMS and an MS and PhD in economics from Yale University.

Dr. Jacob N. Shapiro is Assistant Professor of Politics and Public Affairs at Princeton University and co-Director of the Empirical Studies of Conflict Project, a multi-university research effort between Princeton, Stanford, and UC San Diego. He brings three areas of expertise to the project. The first is his work on the political economy of government service provision. The second is his work on relationships between labour market outcomes and political behavior in a range of countries, work that has included two large nationally-representative surveys in Pakistan (n=6,000 and n=16,000). The third is his experience employing innovative survey methods for eliciting sensitive attitudes and information while minimizing a range of response biases. His research has been published in numerous peer-reviewed economics and political science journals including *Journal of Political Economy*, *American Journal of Political Science*, *International Security*, *International Studies Quarterly*, *Journal of Conflict Resolution, Political Analysis,* and *World Politics* among others. He is a term member of the Council on Foreign Relations, an Associate Editor of *World Politics*, and is a faculty fellow of the Association for Analytic Learning about Islam and Muslim Sociteies (AALIMS). Shapiro received a BA degree in political science from the University of Michigan and a PhD in political science and MA in economics from Stanford University.

Additionally, the following researchers will provide targeted technical gender expertise and oversight of the qualitative field instrument and methodologies:

- Dr. Feyza Bhatti (University of Edinburgh)
- Dr. Rabea Malik (IDEAS)

Individual profiles of these researchers are given below:

Dr. Feyza Bhatti received her BA (Honors) in Economics from Bilkent University, MS in Economics from Eastern Mediterranean University and PhD in Sociology from University of

Edinburgh. She is a mixed methods social science researcher for over a decade with substantive experience and interest in gender and development, sociology of family, public and population health (maternal and child health, nutrition, reproduction and fertility), and disability in South Asian contexts. Between 2002 and 2010, she worked as a Senior Research Fellow and Deputy Director Projects at Mahbub ul Haq Human Development Centre, where she wrote gender-related background papers for the annual Human Development in South Asia reports (2002-2007) and conducted quantitative and qualitative research under the DfID funded five-year multi-country project, Research Consortium in Outcomes of Education and Poverty (RECOUP). For the RECOUP project, she was the lead researcher responsible for the qualitative research activities in Pakistan, and led the Health and Fertility Study, Disability, Education and Poverty Project, and the project on Skill Acquisition and its impact on lives and livelihoods in Punjab and Khyber Pakhtunkhwa.

Dr. Rabea Malik is a research fellow at the Institute of Development and Economic Alternatives Rabea's substantive interests lie in the areas of sociology of education and political economy of education reform in low-income country contexts. With a background in policy analysis using mixed methods (MPhil Ed. University of Cambridge) and research on markets, and inequities in school choice and parental participation (PhD University of Cambridge), Rabea's current research interests include marketization of primary and secondary education; inclusive education; and school based management. She has conducted policy research studies on the education-poverty nexus in low-income country contexts, the political economy of aid for development, and alternative service delivery mechanisms in education. She has also been involved with a multi-year, multi-country research consortium on outcomes of education, where she conducted qualitative research on areas including health and fertility outcomes, and skill acquisition and its impact on lives and livelihoods in Pakistan. She is currently conducting a mixed-methods study on school based management and public private partnerships in education, which is being funded by DfID.

These researchers will be accompanied by a research team which will be composed of the following: a Project Coordinator, three Research Associates, two Research Assistants and three Field Coordinators. Overall, the team will assist the researchers in designing interventions and conducting qualitative and quantitative analysis; and will be responsible for monitoring and management of survey activity and evaluation roll-outs. Other than the Field Coordinators, who are stationed in the Pilot Program Districts, the team will be based at CERP.

Bibliography

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

Ali, K. (1991). Problems of working women in the rural informal sector of Multan district. *South Asian Studies, 8,* 1, 63-80.

Cheema, A., and M. F. Naseer. (2010). Poverty, Mobility and Institutions in Rural Sargodha: Evidence for Social Protection Reform. Report submitted to the Planning and Development Department (P&DD) of the Government of Punjab.

Cheema, A., and M. F. Naseer. (2013). Historical Inequality and Intergenerational Educational Mobility: The Dynamics of Change in Rural Punjab. *The Lahore Journal of Economics*, 18: 211-232.

Amjad, R. (2013). Why Has Pakistan Not Reaped Its Demographic Dividend?.*Population Council Book Series*, *1*(1), 41-53.

Durr-e-Nayab. (2008). Demographic Dividend or Demographic Threat in Pakistan?. *The Pakistan Development Review*, 1-26.

Aslam, M., and Rawal, S. (2013). Preparing Women of Substance? Education, Training, and Labor Market Outcomes for Women in Pakistan. *The Lahore Journal of Economics*, 18: 67-93.

Nasir, Z. M. (2003). Population of Pakistan: An Analysis of 1998 Population and Housing Census. In Kemal, A. R., Irfan, M., & Mahmood, N. (Eds.), *Economically Active Population, Employed and Unemployed; An Evaluation of the 1998 Population Census Data*. Islamabad: Pakistan Institute of Development Economics

Nayar, R., Gottret, P., Mitra, P., Betcherman, G., Lee, Y., Santos, I., Dahal, M., & Shrestha, M. (2012). *More and Better Jobs in South Asia* (1st ed.). World Bank Publications.

Arcury, T. A., Preisser, J. S., Gesler, W. M., & Powers, J. M. (2005). Access to Transportation and Health Care Utilization in a Rural Region. *The Journal of Rural Health, 21,* 1, 31-38.

Attanasio, O., Kugler, A., & Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *3*(3), 188-220.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., & Sulaiman, M. (2012). Empowering adolescent girls: Evidence from a randomized control trial in Uganda. *Unpublished Working Paper*.

Barrera-Osorio, F., Bertrand, M., Linden, L. L., & Perez-Calle, F. (2008). Conditional cash transfers in education design features, peer and sibling effects evidence from a randomized experiment in Colombia (No. w13890). National Bureau of Economic Research.

BenYishay, A., and Mobarak, A. (2013). Communicating with Farmers Through Social Networks. Yale University Economic Growth Center Discussion Paper No. 1030; Yale Economics Department Working Paper No. 121. Available at SSRN: http://ssrn.com/abstract=2315229

Burde, D., & Linden, L. L. (2013). Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools. *American Economic Journal: Applied Economics*, *5*(3), 27-40.

Card, D., Ibarrarán, P., Regalia, F., Rosas-Shady, D., & Soares, Y. (2011). The labor market impacts of youth training in the Dominican Republic. *Journal of Labor Economics*, *29*(2), 267-300.

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

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Sheikh, S., Siddiqui, S., Tourek, G., Emeriau, M. (2012 b). The Employers Survey Report, Center for Economic Research in Pakistan.

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Lodhi, A., Sheikh, S., Siddiqui, S., Tourek, G., Niazi, M., Shoaib, A. (2012 c). The Skills for Employability Evaluation Report, Center for Economic Research in Pakistan.

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Lodhi, A., Sheikh, S., Siddiqui, S. (2013 a). The *SFM*-Village Based Training Evaluation Report, Center for Economic Research in Pakistan.

Cheema, A., Khwaja, A., Naseer, F., Shapiro, J., Lodhi, A., Sheikh, S., Siddiqui, S. (2013 b). The *SFJ*-Stipend Evaluation Report, Center for Economic Research in Pakistan.

Ekirapa-Kiracho, E., Waiswa, P., Rahman, M. H., Makumbi, F., Kiwanuka, N., Okui, O., & Peters, D. H. (2011). Increasing access to institutional deliveries using demand and supply side incentives: early results from a quasi-experimental study. *BMC international health and human rights*, *11*(Suppl 1), S11.

Flora, N. and Paniagua, G. (2013). Meta-Evaluation of Private Sector Interventions in Agribusiness, IFC, Washington.

Gammage, S., Diamond, N., and Melinda P. (2005). Enhancing Women's Access to Markets: An Overview of Donor Programs and Best Practices. U.S. Agency for International Development (USAID).

Glennerster, R and Takavarasha, K. (2010). Empowering Young Women: What do we know? Prepared for the Nike Foundation by The Abdul Latif Jameel Poverty Action Lab at MIT.

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2012). *Learning through noticing: theory and experimental evidence in farming* (No. w18401). National Bureau of Economic Research.

Hicks, J. H., Kremer, M., Mbiti, I., & Miguel, E. (2011). Vocational Education Voucher Delivery and Labor Market Returns: A Randomized Evaluation among Kenyan Youth, Report for Spanish Impact Evaluation Fund (SIEF) Phase II. *Policy Note Human Development Network. World Bank*.

IEG World Bank-IFC-MIGA (2011). Impact Evaluations in Agriculture: An Assessment of the Evidence. World Bank, Washington.

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions. *The Quarterly Journal of Economics*, *126*(1), 145-205.

Leibowitz, A. A., & Taylor, S. L. (2007). Distance to public test sites and HIV testing. *Medical care research and review*, *64*(5), 568-584.

Maitra, P. and Mani, S. (2012). Learning and Earning: Evidence from a Randomized Evaluation in India. Monash University Department of Economics Discussion Paper 44/12.

Malik, S. (2005). Agricultural Growth and Rural Poverty: A Review of the Evidence. ADB Pakistan Resident Mission Working Paper Series. Working Paper No. 2.

MCC. (2012). MCC's First Impact Evaluations: Farmer Training Activities in Five Countries. Millenium Challenge Corporation Issue Brief.

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

Mumtaz, Z., & Salway, S. (2005). 'I never go anywhere': Extricating the links between women's mobility and uptake of reproductive health services in Pakistan. *Social Science and Medicine : an International Journal, 60,* 8, 1751-1765.

Ñopo, H., Robles, M., & Saavedra, C. J. (2007). *Occupational training to reduce gender segregation: The impacts of ProJoven*. Washington, DC: Inter-American Development Bank, Research Dep.

Rashid, R., and Vigoda, M. (1990). Women in Local Markets and Commercial Access: A Report and Handbook. Dhaka, Bangladesh: USAID.

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

Rawlings, L. B., & Rubio, G. M. (2005). Evaluating the impact of conditional cash transfer programs. *The World Bank Research Observer*, *20*(1), 29-55.

Sathar, Z. A., & Kazi, S. (2000). Women's autonomy in the context of rural Pakistan. *The Pakistan Development Review*, 89-110.

Thornton, R. L. (2008). The demand for, and impact of, learning HIV status. *The American Economic Review*, *98*(5), 1829.

Appendix A: Proposed Outputs and Corresponding Budget Payments

Total Budget: GBP 1048,914

Output	Description	Expected Date of Submission	% of Budget	Amount (GBP)
Report on Design and Compliance of Uptake Interventions in SFM 2013-14	This report will detail the motivation, objectives, design of interventions, and evaluation design of calliberations to PSDF's SFM 2013-14 scheme aimed at increasing access to training for rural women. It will also report on the sample for the evaluation as well as compliance on the different interventions using tracker surveys 1 & 2.	Nov-14	20%	209,783
Report on Uptake and Maximizing participation for marginalized women	Evaluation of design calibration in SFM 2013-14 intervention. This report will provide details on interventions carried out as part of design calliberations to SFM 2013-14 scheme and give details on what impact each intervention had on take-up of SFM 2013-14 training.	Dec-14	20%	209,783
Report on Design of Market Linkage Intervention	Report on the motivation, objectives, evaluation design and design of a market linkage intervention which will be implemented with graduates of PSDF's SFM 2013-14 scheme. The document will also provide findings from baseline evidence on the evaluation sample as well as details on the sampling methodology for the evaluation.	Jan-15	20%	209,783
Report on Design of Job Matching Intervention	Strategy document on the motivation, objectives, evaluation design for evaluating a job placement intervention as design calliberations to PSDF's Skills for Jobs (SFJ) scheme. Using baseline evidence, this report will detail on the design of the intervention itself as well as the sampling methodology for the evaluation.	Feb-15	25%	262,228
Report on Design of BPSV Scheme	Design concept of the *Big Push on Skills for Villages* (BPSV) scheme using in-depth baseline evidence and analysis. The document will detail the motivation behind the intervention and evaluation, it's objectives, and evaluation methodology for evaluating the intervention. This report will also provide details on the intervention design and the sampling methodogloy for conducting the evaluation.	May-15	15%	157,337

Appendix B: Extract from the Minutes of 17th Meeting of the PSDF Board

Extract from the Minutes of 17th Meeting of the PSDF Board held on the 7th of February 2014:

viii. Evaluation Methodology

The Board approved the following evaluations as part of the PSDF-CERP collaboration:

a. The evaluation of the skills for market (SFM) scheme for rural women in the pilot districts for measuring economic and non-economic returns to the training for rural women with and without market linkages. In addition, the evaluation will estimate the impact of village-based training, transport, stipends and mobilization on training uptake.

b. The Board reaffirmed its approval, given in the 16th Board meeting, of the evaluation of a big push scheme that provides village-based training in agriculture and related skills in the 4 pilot districts (Bahawalpur, Bahawalnagar, Lodhran and Muzzafargarh) and requested CERP to include details in its proposal.

c. The Board also approved the evaluation of a sample of the core PSDF schemes in the expansion districts using oversubscription design methodology.

d. The Board also approved the evaluation of a pilot intervention that complements the Skills for Jobs (SFJ) scheme with interventions designed to strengthen job placement. It approved a joint PSDF-CERP Committee (with Dr. Ijaz Nabi,Chairman and Mr. Almas Haider, Member from the PSDF Board) to finalise the design of the intervention to be evaluated.

Appendix C: Status of the *SFM* 2013-14 Evaluation

The design of interventions to mitigate access constraints was finalized in agreement with PSDF in November 2013. Randomized offers of training to one female household member were made to approximately 8,200 (treatment) CERP sample households in 285 CERP baseline sample villages as part of the evaluation during December 2013 and February 2014. This offer was for enrollment in the tailoring course for women being offered as part of *SFM* 2013-14 scheme. As part of the evaluation the location of the 150 training centers was randomized in order to estimate the additional effect of locating a training centre in a village. The randomization has ensured that centers are located in all *tehsils* in three program districts (Muzzaffargarh, Bahawalpur and Bahawalnagar) and this has broadened geographical access. Offers of interventions (see section 3) designed to mitigate access constraints have been made to random sub-samples of the treatment population. Enrollment in all the evaluation centers was finalized in March 2014 and training has begun in all centers. Initial results show an extremely large increase in applications and uptake compared to previous *SFM* rounds. There are 8,408 applications for 3,000 slots and an uptake of nearly 40% in CERP voucher holders, which is an increase of 35 percentage points over SFE. The fact that there has been exceptional demand, even the remote villages of the districts, compared to previous rounds shows that the interventions were well-designed. It also shows that the low uptake revealed as part of the *SFE* evaluation really reflected underlying access constraints. The evaluation will provide estimates about the relative cost-effectiveness of different interventions in mitigating access constraints and will also provide results on economic and non-economic returns to *SFM* training.

Appendix D: Evidence and Program Impact from the Baseline Phase

CERP has generated rich and extensive context-specific evidence of great relevance to the design of the PEOP program using two sources:

a. Large-scale baseline surveys of household and employers in the program districts. This has involved *tehsil*-representative surveys of approximately 32,000 households in 861 communities (581 villages and 280 urban neighborhoods), in-depth surveys and censuses in 149 villages and representative surveys of 6000 employers and 3,500 livestock and diary suppliers in program districts.

b. Phased evaluations of early stage interventions designed to increase training uptake among rural women as part of the SFM scheme and among urban males as part of the SFJ scheme. In addition, a phased evaluation of the SFE scheme was conducted to understand the variations in training uptake within the target population and to provide evidence on the correlates of low uptake. An initial evaluation has also been conducted of the farmer's day scheme implemented by the Government of Punjab's Livestock and Dairy Development Department. These evaluations use the RCT methodology and are in line with the global gold standard. The evaluations form the basis of the model of smart policy design that was evolved during the baseline phase.

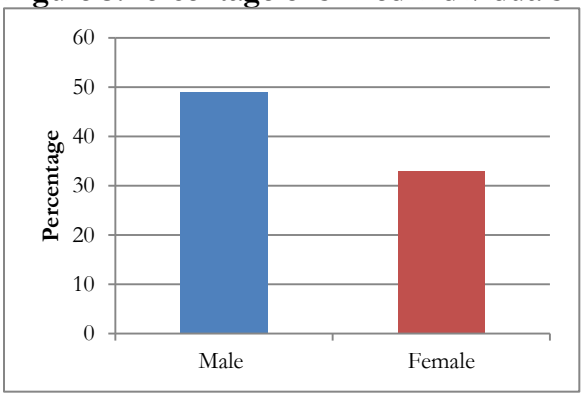
CERP's activities in the baseline phase of the project have focused primarily on creating evidence, learning and small-scale experimentation to prepare for larger-scale impact evaluations and to enable a program design that has the potential to demonstrate high impact. In addition, the evidence generated during the baseline phase had provided a baseline for log frame monitoring and multiple evaluations of PSDF interventions. The evidence has also been used by PSDF to program its menu of trades. Finally, the evidence has also provided a rich set of stylized facts (skills demand and expectation of returns, employment and income expectations from training, job search networks, skills gaps facing employers, state engagement and poverty, distribution of livestock asset holding, access to dairy markets, organizational structure of dairy markets etc.[35]) and insightful analysis about supply and demand challenges in the skills, labor, livestock and dairy markets and an understanding of expectations, constraints and behavior among households and employers in these markets.

Specifically, several key, context-specific insights have emerged from the evidence produced in the baseline phase and for skills these can be grouped into the following eight areas of first-order design challenges that have had direct program impact:

1. ***Paucity of Training Supply in the Program Districts:*** Baseline surveys provide rigorous evidence about ***the existence of skills deficits in the program districts*** (Figure 3) and ***the paucity of training supply in the region*** (Figure 4)**.** This evidence was used by PSDF to corroborate its initial direction that prioritized the need to setup an effective model of seeding training supply in the pilot districts.

[35] The details of this analysis can be found in the six reports produced by CERP from the baseline data, which have been made available to PSDF, Government of Punjab and DFID.

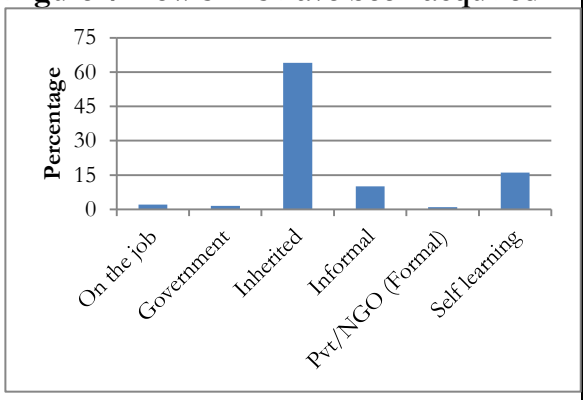
Figure 3.Percentage of skilled individuals



Source: CERP Baseline Household Survey

Note: Self-reported by male and female respondents

Figure 4 How skills have been acquired

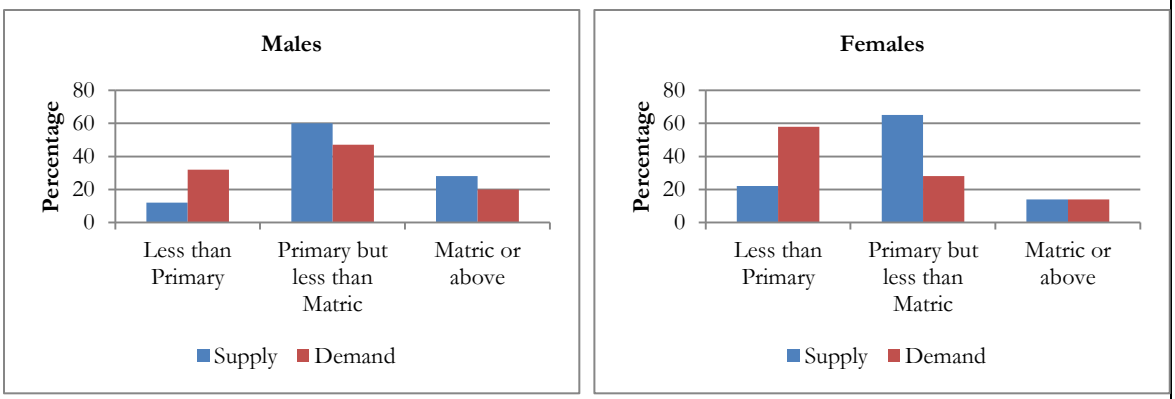


Source: CERP Baseline Household Survey

Figure 3 illustrates the deficit of skilled males and females in the general population of the pilot districts. Figure 4 shows that before PSDF got established, a majority of individuals (approximately 76%) did not acquire skills through formal private or government skills programs.

2. ***Existence of Demand-Supply Mismatches:*** Baseline surveys provide rigorous evidence about the ***existence of supply-demand mismatches*** (Figure 5) and ***the need to seed providers that supply training programs that match the skills needs of the target population and use offerings and pedagogy that are accessible to a population with low educational attainment.*** This evidence was used by the PSDF Board to broaden its menu of trades and program the menu to make it attractive and accessible to the target population whose members have low levels of education attainment. This was done through the introduction of the SFM scheme. In addition, basic numeracy and literacy have been added as integral components of the training courses in this scheme.

Figure 5 Demand - supply mismatch in education requirement



Source: CERP Baseline Household Survey (Demand) and PSDF (Supply)

Note: Education requirements for training are for courses offered in PSDF's *Skills for Employability* (SFE) scheme. Percentages are calculated as a fraction of total courses offered for each gender

Figure 5 illustrates an instance of demand-supply mismatches in the provision of vocational training in the form of a mismatch between educational attainment in the general population and educational requirements set by training providers in the initial PSDF schemes This shows that the mismatch was most acute for the less educated who form the majority population in the district and among the poor.

3. ***Low Mobility and Dependence on the Local Market:*** Baseline surveys provide rigorous evidence that ***the target population, especially women, has low levels of mobility and is focused on the local labor market*** (Figure 6) and therefore ***effective program design will need to bring training to the doorstep and offer training menus that have the potential to augment incomes in the local market.*** This evidence was used by the PSDF Board to introduce the SFM and the SFF schemes and to program training locations in a way that mitigates the urban bias in location and thereby improves rural citizens' access to training.

Figure 6 Location of work

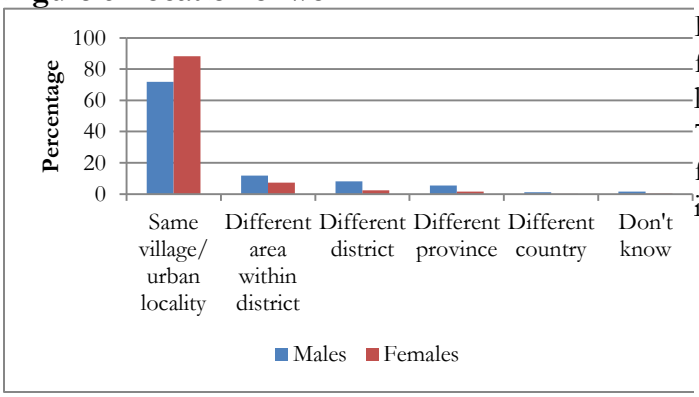
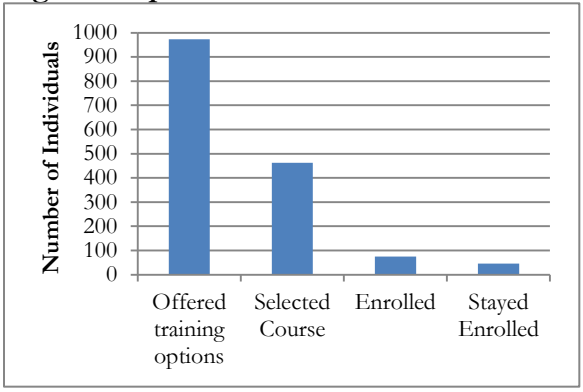


Figure 6 shows that both, males and females, are focused on the local labor market with approximately 70% and 85% of the male and female working population working in the same village/ urban locality.

Source: CERP Baseline Household Survey

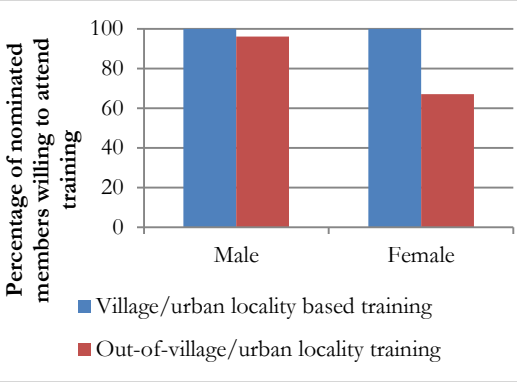
4. ***Low Uptake for PSDF Training:*** The phased evaluation of the SFE shows that: (a) ***program uptake is low in the target population*** (Figure 7) and (b) ***low uptake does not reflect low demand but is a consequence of real constraints related to distance and mobility and the expectations of getting jobs and augmenting income after acquiring*** (Figure 8). This evidence was used by the PSDF management to design interventions to: (a) increase uptake among rural women by locating training centers in villages as part of the *SFM* scheme and (b) increase uptake among urban males by increasing stipends as part of the SFJ. These interventions have been evaluated through phased evaluations by CERP and the PSDF program has been calibrated in response to the evidence produced by this set of evaluations. In addition, PSDF is designing models that integrate job placement and market linkages with skills provision, which will not only address challenges related to uptake but also increase the value-addition from training.

Figure 7 Uptake in SFE



Source: Phased Evaluation of PSDF's SFE scheme

Figure 8 Preference for training location



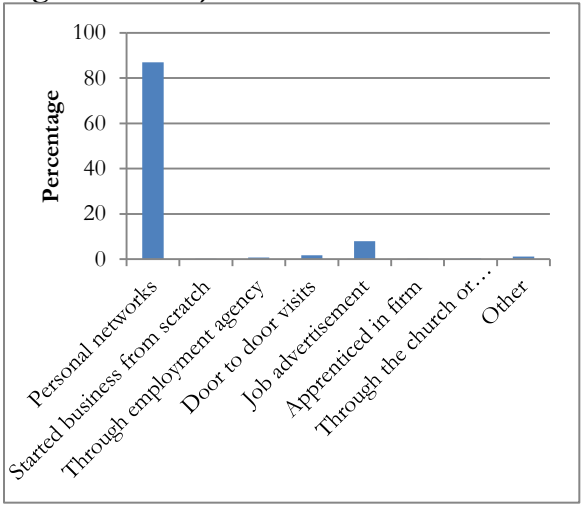
Source: CERP Baseline Household Survey
Note: Calculated as a percentage of male and female members who were: (i) nominated by household heads for vocational training during the Baseline Household Survey; and (ii) willing to attend training

Roll out of the RCT based phased evaluation of PSDF's *Skills for Employability* (SFE) scheme consisted of a 3-stage process with the following steps (more detailed provided in section 5): (i) training offered through vouchers; (ii) selection of course being offered in SFE; and (iii) enrollment in course being offered in SFE. Figure 7 shows the uptake across each stage of the roll out of the phased evaluation of SFE. It shows that out of the 973 individuals who were offered training options, only 75 individuals enrolled in training courses; and more than 25 out of these 75 dropped out.

Figure 8 shows that the willingness to enroll in training among women decreases acutely with distance; i.e. distance is a constraint towards realizing the demand for training.

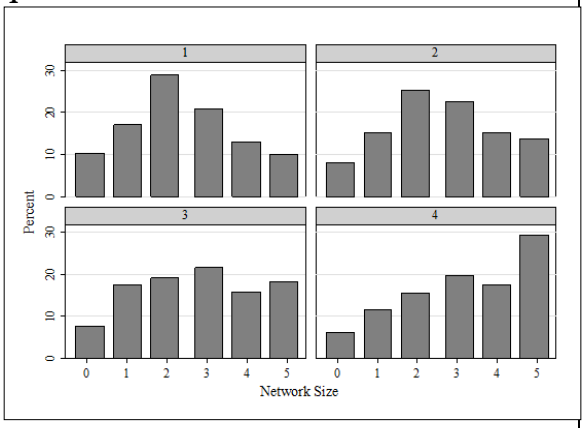
5. ***Personalized and Exclusive Job Search and Placement Networks:*** Baseline surveys provide rigorous evidence that ***job placement and search networks are highly personalized and exclusive*** (Figure 9 and 10) and that ***both the uptake*** and the ***value addition of training is likely to remain low unless effective job placement is integrated with skills training.***

Figure 9 How job is found



Source: CERP Baseline Household Survey

Figure 10 Network size by consumption quartile



Source: CERP Baseline Household Survey
Note: An index on the scale of 0 – 5 is used to depict network size; where 0 represents narrowest network and 5 represents most diverse network size

Figure 9 illustrates that a majority of the population in the pilot districts find jobs through informal personalized networks. The Baseline Household Surveys also reveal a positive correlation between household consumption and network size (Figure 10); which implies that wealthier households are better placed to take advantage of personalized networks to find jobs.

6. ***Low Uptake in Male Income Earners that is Moderately Responsive to Increases in Stipends:*** The phased evaluation of the SFJ scheme shows that ***in spite of high demand for skills training uptake remains low among the male income earners in the target population*** (Figure) and that ***low uptake persists in this population even after the stipend has been increased 2-3 fold*** (Figure 2)**.** Evidence from this set of phased evaluations along with the evidence on job placement networks has been used by PSDF to enter into a contractual arrangement with specialized job placement providers and it is in the process of designing a cost effective high impact model for job placement. This evidence was also used by PSDF to calibrate stipend amounts.

Figure 11 Uptake by employment status

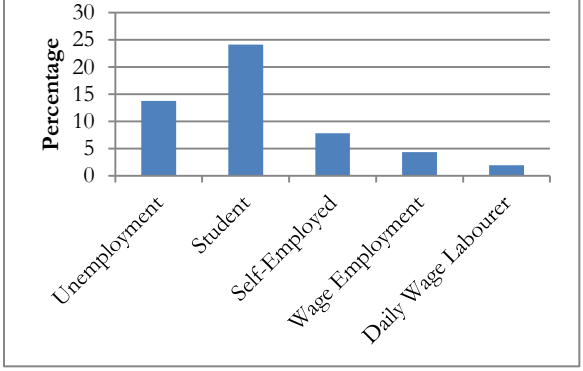
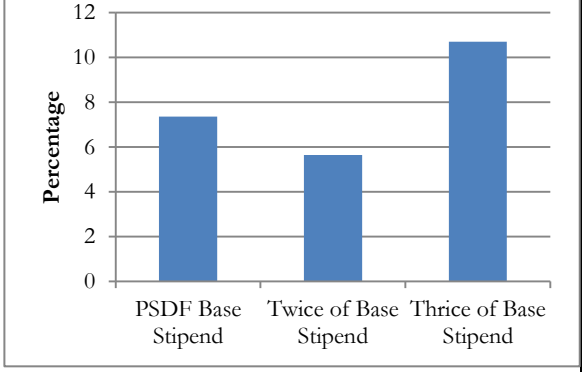


Figure 12 Uptake by stipend level



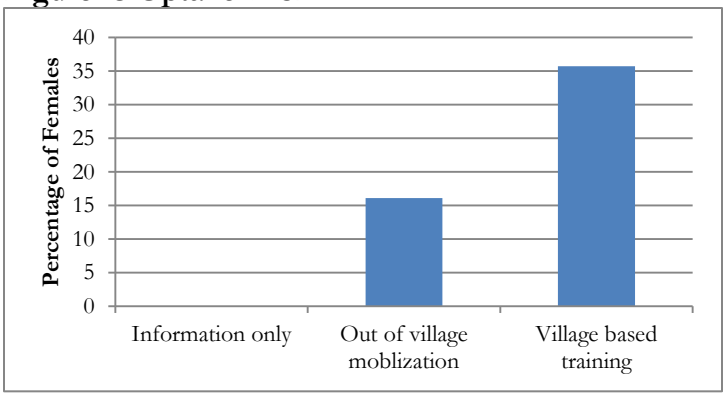
Source: Phased Evaluation of PSDF's SFJ scheme

Source: Phased Evaluation of PSDF's SFJ scheme

As part of the phased evaluation of PSDF's *Skills for Job* (SFJ) scheme, respondents were provided differential stipend to evaluate the impact of increased stipend on uptake of PSDF trainings among urban males. Figure 11shows that uptake remained low among active income earners and was highest for students and unemployed individuals. Figure 12 presents uptake across each stipend level that was offered in the evaluation and shows that uptake remained moderate even when the stipend level offered by PSDF was increased by 2-3 fold.

7. ***Distance Constraints and Low Uptake among Women:*** The phased evaluation of the SFE scheme shows that ***training uptake among women is adversely affected by distance*** (Figure 13) ***and the opportunity cost associated with household level work***. PSDF responded to this evidence by designing an intervention that brought training centers to the village and reduced the distance to the center for potential women trainees. This intervention was implemented as part of the SFM scheme rolled out at the end of 2012. The phased evaluation showed that reducing distance results in a 30 percentage point increases in uptake among rural women (Figure 4). It also shows that the intervention results in positive selection affects with the poor and vulnerable being more likely to access training (Figure). This evidence has resulted in village-based training being mainstreamed as part of the SFM scheme for women.

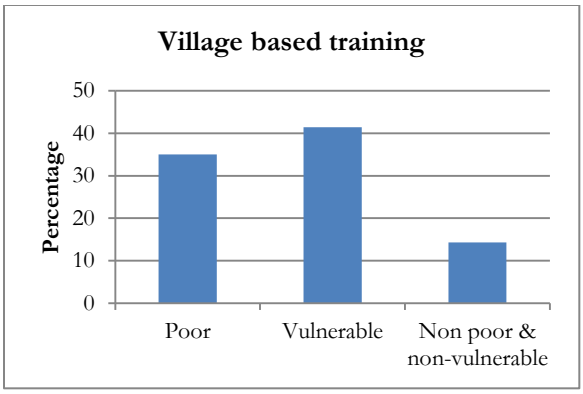
Figure 13 Uptake in SFM

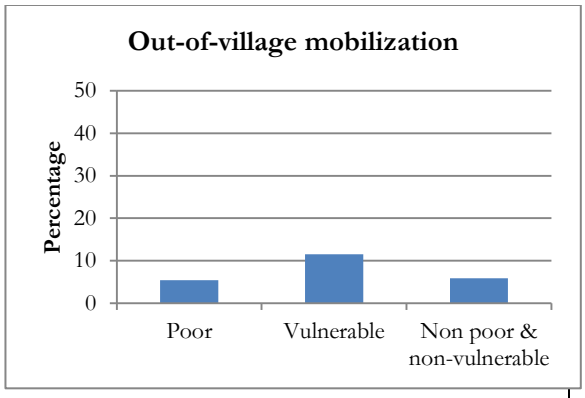


Source: Phased Evaluation of PSDF's SFM scheme

The phased evaluation of PSDF's *Skills for Market* (SFM) scheme evaluated the impact of: (i) village based training; and (ii) social mobilization to encourage people to enroll in out-of-village training, on uptake of training among rural females. Figure 13 shows the uptake across the different interventions under the SFM phased evaluation. This evaluation revealed that reducing distance results in 35 percentage point increase in uptake.

Figure 14 Uptake by poverty status and distance from training





Source: Phased Evaluation of PSDF's SFM scheme

Figure 14 shows that providing village based training as part of SFM phased evaluation had positive selection effects, that is the poor and vulnerable enroll in much higher proportions

8. ***Low Mobility and Poor Market Linkages Among Women:*** Baseline surveys and the phased evaluations show that ***women's labor market opportunities are confined to low skill jobs in the local market and their limited mobility*** (Figure 5) ***restricts self-employed women from effectively linking with markets*** and therefore ***the value-addition of training for women is likely to remain low unless an effective model of market linkages is integrated with skills training.*** PSDF is working with a select set of training providers to develop a cost-effective and high impact market linkage model for women that integrates linkages with skills training.

Figure 15 Employment status of working females

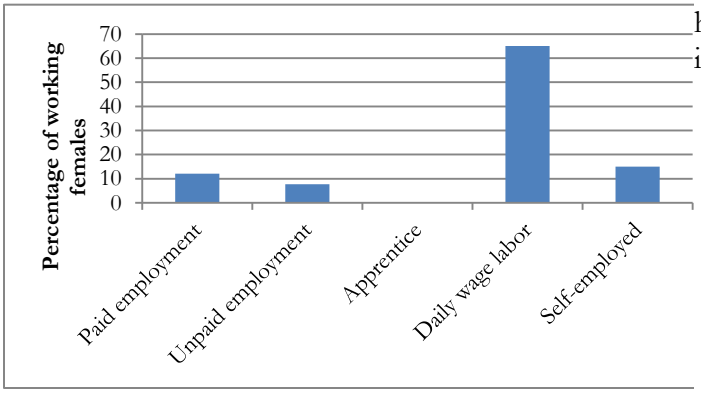


Figure 15 shows that an extremely high proportion of working women in low skilled daily wage work.

Source: CERP Baseline Household Survey

During the baseline phase CERP has addressed two of the above first-order issues through the execution of small-scale phased evaluations in PEOP districts. For example, pilot experiments in PSDF's *Skills for Market* (focusing on rural women) and *Skills for Jobs* (focusing on urban men) schemes employed interventions such as village-based training centers, social mobilization, and differential stipends to test their impact on uptake rates. Complementary job search and placement interventions are currently being designed to address highly personalized job placement networks and preferences for on-the-job training. Finally, *a village-based version of Skills for Farms scheme* is currently being designed that aims to

move the productivity frontier in an entire community through village-based training. The expectation is that this scheme will have the potential to create large-scale community level impact on village-level productivity and income and the earning of members of the target population.

A more complete list of key findings from the baseline household, employer and livestock and diary surveys is given in the detailed reports that have been submitted to PSDF, Government of Punjab and PSDF. Rich results have also been produced on the livestock and dairy markets that are not reproduced here as this component has been discontinued and these results can be found in the relevant reports submitted to DFID. Findings from the baseline surveys and phased experiments will continue to feed directly into future collaborative activities between CERP and PSDF, including large-scale evaluative RCTs.

Appendix E: Protocols for Data Quality Assurance

To ensure quality of data, both at the collection and at the entry level, a number of mechanisms have been designed by CERP that guarantee data accuracy and precision. Though a survey company is contracted to carry out the survey processes, all the procedures are thoroughly drafted and scrutinized by the CERP team. The following methods have been put in place for data quality assurance: from the initial questionnaire design phase till the point of data entry:

- Designing of the questionnaires and survey instruments with pretesting
- Training of enumerators and the development of the survey manual
- Data collection and conducting field surveys – survey team structure and quality monitoring
- Data entry

Separate instruments are designed to capture qualitative and quantitative data. Survey instruments take into account regional/local characteristics that can impact deliverables, of which language is of utmost importance. CERP designs questionnaires in English and forwards it to the survey company for translation into the local language (mostly Urdu and Punjabi). The translated versions are then back-translated by the CERP research team to ensure that true meaning of the questionnaires is preserved. Then the most suitable translation is finalized for administering in the field. These initial drafts are then piloted in the presence of the CERP research team at the targeted areas to ensure that the required data is being captured. Moreover, qualitative pilot interviews are also recorded to check for their suitability. Questionnaires are then remodeled based on feedback received from this pretesting.

In-depth training of enumerators and their supervisors is conducted to coach them about survey procedures, quality standards, ethics and responsibilities. A member of the CERP research team is present during these training sessions to ensure that proper training is being delivered. To enhance enumerator understanding, the training also includes briefings on survey goals, survey modules and survey questions. Pilot surveying is carried out by enumerators which is back-checked with on-spot feedback/suggestions by the trainer. More than the required number of enumerators are called for this training and a screening process is carried out to select the final team of enumerators based on their performance during the training. Furthermore, a survey training manual is constructed to guide enumerators through the process. This survey manual comprises of techniques that should be applied for data capturing. It also states the questions in detail to clear any ambiguities in the survey. Thus, it helps maintain consistency in data collection by enumerators across regions and time.

A survey team structure is in place to monitor the quality of data collected. A supervisor per 4-5 enumerators is recruited to oversee the enumerators' work. The supervisor checks questionnaires filled enumerators and channels them back to respondents if the required information is incomplete or incorrect. He also manages enumerators' schedule and work progress. On top of this, field editors working at the main area base ensure surveys are filled correctly before they can be sent for data entry. Data entry is carried out by the survey company near the field so that potential errors in data collection can be minimized. At least

10% randomly assigned interviews are conducted in the presence of monitors. Back-checking or revisits are also be carried out by these monitors to ensure data precision. Every survey is rechecked by a team of scrutinizers for errors in data entry. CERP also assigns Research Assistants/Associates to conduct systematic random checks on field activity timelines, monitoring activities and also oversee error reports to ensure data quality.

Double entry of data is carried out using a software package by the survey company, while CERP research staff randomly enters some of the data to check for errors as well. This helps ensure that data provided by the survey company is below the agreed error rate (i.e. 0.5%).

CERP team stays in constant contact with the survey company throughout the survey process, to make sure that the set protocols are being followed. Progress reports on data collected and enumerator performance are reviewed by CERP team to check for any discrepancies, error patterns and the flow of work.