HCEO WORKING PAPER SERIES

Vocational Training Programs and Youth Labor Market Outcomes: Evidence from Nepal

Shubha Chakravarty
Mattias Lundberg
Plamen Nikolov
Juliane Zenker

Working Paper

2017-056

07/2017

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Vocational Training Programs and Youth Labor Market Outcomes: Evidence from Nepal*

SHUBHA CHAKRAVARTY[1], MATTIAS LUNDBERG[2,], PLAMEN NIKOLOV[2,3], and JULIANE ZENKER[4]

Lack of skills is arguably one of the most important determinants for high levels of unemployment and poverty. Targeting youth unemployment and also important because of its strong influence on other important social outcomes. Using a "fuzzy" regression discontinuity design, we examine the employment effects of a vocational training program in Nepal launched in 2009 over a three-year period. We find program participation generated an increase in non-farm employment of 28 percentage points for an overall gain of 95 percent, three years into the program. The program also generated an average monthly earnings gain of 2,167NRs (≈ 29 USD) or 171 percent. Applying heterogeneous local average treatment effect (HLATE) estimators, we find striking differences in the impacts by gender: program impacts are almost double the size for women than for men.

Keywords: Training, Employment, Labor, Economic Development, Nepal, Regression Discontinuity, Gender

JEL Codes: I21, I28, I38, J08, J24, O15

---

*We thank Siroco Messerli and Bal Ram Paudel for facilitating all aspects of this study. Data collection for this paper has been supported through the Swiss Development Corporation (SDC), the UK's Department for International Development (DFID), The World Bank's Multi-donor Trust Fund for Adolescent Girls. Headed by Madhup Dhungana, New Era Limited provided exceptional support with survey design, survey implementation and data management. At the World Bank, Jasmine Rajbhandary, Venkatesh Sundararaman, and Bhuvan Bhatnagar. Amita Kulkarni, Uttam Sharma and Jaya Krishna Upadhaya coordinated survey activities. Ali Ahmed, Marine Gassier, and Jennifer Heintz provided outstanding research support.

[1] The World Bank
[2] State University of New York (at Binghamton), Department of Economics. Email: nikolov.pv@gmail.com
[3] Harvard University Institute for Quantitative Social Science
[4] University of Göttingen. Email: jzenker@uni-goettingen.de

I. Introduction

Lack of skills is arguably one of the most important determinants for high levels of unemployment and poverty (Carneiro and Heckman 2003; Heckman and Krueger 2004; Holzer 2007; Malamud and Cristian Pop-Eleches 2010). The nexus between skill formation and employment is particularly relevant for youth in the labor market. In much of the developing world, unemployment among the youth is extremely high: 17 percent of the world's population are youth (ages 16-24) and youth make up 40 percent of the world's unemployed (United Nations 2012; World Bank 2017). Targeting youth unemployment and underemployment is important also because of its strong influence on other important social outcomes: unemployment negatively impacts crime rates (Blattman and Annan 2016; Fella and Gallipoli 2007), depression prevalence (Frese and Mohr 1987), substance abuse rates (Linn, Sandifer and Stein 1985), and rates of social exclusion (Goldsmith, Veum and Darity 1997). Therefore, targeting unemployment with effective interventions is one of the highest priorities in low-income countries (World Bank 2013). One common policy response in effort to enhance skill formation among the youth is vocational training programs. To this end, the European Union launched in 2013 an eight billion euro initiative aiming to provide every young European with a job, apprenticeship, or training within four months of becoming unemployed. Similarly, in Latin America, job training programs (referred to collectively as the "*Jovenes*" programs) have been implemented since the early 2000s. To date, around 700 youth employment programs from 100 countries have been implemented and more than 80 percent of them offer skills training.[5]

In this paper, we examine one of the largest youth training interventions in Nepal, serving almost 15,000 youth. A component of the training program, called the Adolescent Girls Employment Initiative (AGEI) and launched in 2009, specifically targeted young Nepali women. The initiative served approximately 4,500 young women over a period of three years. Program eligibility was based on a discrete threshold in a continuous individual score, which was determined by training providers. Therefore, we exploit the individual-assigned score to implement a regression discontinuity design and estimate the short-term effects of the skills training and employment placement services in Nepal. Because we find some evidence that the

[5] See Youth Employment Inventory (http://www.youth-employment-inventory.org/)

actual individual assigned score was manipulated in practice, we instead use survey data to reconstruct the underlying score components and generate our own individual-based score. We then use the reconstructed score as instrument for training eligibility. We find, approximately twelve months after the start of the training program, the intervention generated an increase in non-farm employment of 0.28 percentage points for an overall gain of 93 percent. The program also generated an average monthly earnings gain of 2169 NRs monthly (≈28 USD). Given that the average monthly income at baseline was 1272 NRs (≈ 17 USD), the program impact resulted in income gain of about 71 percent for the combined 2010-2012 program cohorts. Following Becker et al. (2013)'s approach for heterogeneous local average treatment effects (HLATE) in the context regression discontinuity set-up, we estimate heterogeneous treatment effects by gender and type of trade. We find that the program's impacts on employment were larger for women than for men. We also detect gender differences for trade-specific employment, earnings, and gainful employment.

Despite the rapid expansion of skill-enhancement employment programs across the world, debates about their positive impacts on employment outcomes still persist (LaLonde 1995; Heckman, LaLonde and Smith 1999).[6] Based on US and European evidence, Card et al. (2009) review impacts of various training programs. They suggest that youth programs tend to yield less positive impacts than untargeted programs and on-the-job training programs are not particularly effective in the short run but have larger positive impacts after two years. In contrast, our study detects large positive impacts on employment and earnings even in the short-run. Interventions from developing countries in general show larger impacts than programs conducted in other regions.[7] Participants in comprehensive training programs have higher probability of finding a job than single component programs and the program this study examines, participants participate in a single training workshop that leads to large subsequent employment impacts. Estimates of program impacts from developing countries are largely based from interventions in Latin American (e.g., Attanasio, Kugler, & Meghir 2011; Card et al. 2011). Yet, for the most part, these

[6] See Heckman (1999) and many others, including randomized controlled trials, like the Job Training Partnership Act in the US.
[7] See LaLonde (1986), Card and Sullivan (1988), Burghardt and Schochet (2001), Betcherman, Olivas and Dar (2004), Elías et al. (2004) Card et al. (2007), Chong and Galdo (2006). Based on 289 youth employment interventions in 84 countries, Betcherman et al. (2007) show higher impact in developing countries than in developed ones. Most of the rigorous evidence on training programs in developing countries is from Latin America, where positive impacts are particularly pronounced (Gonzalez-Velosa et al. 2012; Attanasio et al. 2008; Reis 2015). Attanasio et al. (2008) evaluate the *Jovenes en Acción* job training program in Colombia. *Jovenes en Acción* provided three months of classroom training followed by a three month unpaid internship at a company. Attanasio et al. (2008) detect positive employment effects for women (4 to 7 percentage points), no employment effects for men and positive earnings effects for both men (8 percent) and women (18 percent). The study argues that the increase in earnings is due to increased employment in formal sector jobs upon training completion.

studies provide skepticism regarding the cost-effectiveness of training programs (Almeida et al., 2012).

Our study contributes to the existing literature on training programs in three distinct ways. First, we present striking results based on a low-income country context and specifically focused on the South Asia region, the world's most populated region which the literature has overlooked despite the region's high youth unemployment rates. Although numerous studies focus on the impacts of training programs in high-income countries, only one other quasi-experimental study in South Asia examine the impact of training programs on employment. Maitra and Mani (2017) evaluate a training program, which bundled a unique feature introduced to increase commitment and encourage regular attendance, in stitching and tailoring offered to young women in poor slum communities of New Delhi. The study found that the program increased the likelihood of casual or permanent wage employment by more than 5 percentage points, self-employment by almost 4 percentage points, and any employment by 6 percentage points. The program increased hours worked in the post-training period by around 2.5 hours.

Second, our quasi-experimental analysis shows large positive results in the short-run that are higher than estimates from recent experimental interventions in developing countries. Attanasio et al. (2011) and Alzua et al. (2013) show respectively gains of 14 percentage points and 10 percentage points. Card et al. (2011) on a program in the Dominican Republic, which also finds positive, though insignificant, effects on earnings. In terms of earnings, this intervention generated substantially higher impacts than the increase in earnings of 18 percent found by Attanasio et al. (2011). Finally, a related BRAC field experimental project in Uganda implemented village-level girls' clubs to provide life skills, reproductive health, and livelihood skills to young women aged 14 to 20. This experimental intervention detected substantial increases in employment (72 percent) (Bandiera 2012).

Third, we report striking positive labor market improvements for the female sample. Females in our sample exhibit higher returns to training as compared to the males. In contrast, the impacts by gender in previous studies is mixed. Kluve (2006) reviewed a number of interventions in Europe and found that the programs had larger impacts for women than for men. However, Card (2009) concludes that there is no evidence that training programs have differential impacts for men versus women.

Section II details the Employment Fund training program in Nepal and the design of the intervention. Section III outlines our empirical strategy. Section IV presents our results. Section V provides various robustness checks. Section VI concludes.

II. Study Design and Data Collection

A. Background on The Employment Fund Training Program in Nepal

Started in 2008, the Employment Fund (EF), now one of the largest skills training programs in the country, provides vocational training and placement services under a unique governance structure. Each year, the Employment Fund provides training programs with various training providers. Table 1 provides the total number of training providers, number of training events, and number of trainees. Table 1 shows an increase of total number of program beneficiaries between 2010 and 2012.

Training courses in technical skills vary across a wide range of trades (*e.g.,* incense stick rolling, carpentry, tailoring, welding and masonry). All females receive 40 hours of life skills training (beginning in 2011) and a sub-set of trainees receive a short course in basic business skills. In addition, each trainee is encouraged to complete a skills certification test offered by the National Skills Testing Board (NSTB).

Upon completion of the classroom-based training, the EF places emphasis on job placement services. EF verifies trainees' employment status three months and six months after the completion of the training.[8] Upon verification, training providers receive an outcome-based payment from the EF that is higher for trainees who are employed. The outcome-based payment system creates strong incentives for the training providers to provide placement assistance and provides graduates with an opportunity to put their new skills to work immediately after the training. The training providers receive fixed payments for each individual they enroll and some of the individual characteristics influence future employment prospects, thereby creating incentives for training providers to enroll some applicants. Second, eligibility for training prioritized certain ethnic and disadvantaged groups, therefore creating additional provider incentives for enrollment. The EF provides relatively higher rewards if trainees are placed into "gainful" employment in which they

[8] The employment status of a sample of graduates is verified by EF field monitors three to six months after the completion of the training event.

earn a minimum of 3,000 NRs (≈40 USD) per month. Further, higher rewards are paid if a placed trainee belongs to a vulnerable group.[9]

In 2010, the EF partnered specifically targeted young women aged 16 to 24 (i.e., AGEI population). Training under this Adolescent Girls Employment Initiative (AGEI) proceeded in the same way as it did for other EF trainees, except that certain events had been flagged in advance as likely to attract female trainees. In addition to training course advertisement, the EF sponsored radio and newspaper ads specifically geared towards young women. Many of these ads specifically encouraged women to sign up for non-traditional trades for women, such as mobile phone repair, electronics, or construction.

B. Data and Sample Description

Our primary sources of data are program administrative data and surveys covering three consecutive cohorts of EF trainees (from 2010 to 2012), with two rounds of data collection for each cohort. [10] Figures 1 shows the survey timeline and Figures 2 and 3 depict the study areas.

[Figure 1 about here]

We sample at the training event and the applicant level. The main sampling frame for this study consisted of all training courses sponsored in a given year. The number of training events comprising the sample frame ranges from 598 (in 2010) to 711 (in 2012). Table 1 reports the number of events and participants each year.

[Table 1 about here]

Sampling into this study included a combination of stratified, random and convenience sampling and was done in two consecutive steps. In the first step, we selected training courses and in the second step we selected individuals, both according to standardized procedures detailed below. To select training course we, first, selected the subset of training events from the universe

[9] The definition of "gainful" employment was increased in 2012 to 4,600 NRs (≈60 USD). Throughout this paper, we use the prevailing exchange rate during 2010 and 2011 of 75 NRS to 1 USD.

[10] For the 2010 cohort, a second follow-up was conducted on half of the cohort.

of EF funded training that occurred between January through April of each year.[11,12,13] Second, from the training events offered during these four months, we randomly selected up to 15 districts. Third, from the list of training events that took place in these districts, we randomly selected 20 percent. Table 2 details the resultant sample for the three cohorts. The 2010 event sample comprised 64 events across 30 districts. The 2011 sample comprised 182 events, of which 113 events were dropped from the baseline survey, either because the survey team could not reach the event on the day of applicant selection or because the event was not "oversubscribed".[14] The remaining 69 events in 34 districts were included in the 2011 baseline sample. In 2012, 85 out of 112 sampled events were included in the study sample.

[Table 2 about here]

To sample applicants, a survey team visited each sampled training event on the day when applicant selection happened. Training assignment as well as selection into the studied sample was based on a standardized interview procedure, which assigned scores to applicants in five different categories and added them up to form a total score. We describe this procedure in more detail belwo. Each event's ranking sheet listed the shortlisted applicants from the top-scorer to the bottom one and indicated the threshold (*i.e.,* minimum score) for admission to the course. The here studied individuals comprise of a subset of ranked individuals -- those who fell in the range of 20 % below or above the threshold.

The sampling procedures resulted in a study sample of 4677 across all three cohorts. For the pooled sample (*i.e.,* 2010-2012 cohorts), the study population is 64 percent female and on average 24.5 years old. Fifty-eight percent are married while 51 percent have at least one child. Approximately 59 percent of the sample engaged in some income-generating activity in the month prior to the survey. When we restrict to non-farm income-generating activities, the employment rate falls to 27 percent. At baseline, the average earnings of the pooled sample were 1272 NRs per

[11] Because of the AGEI focus of this study, we prioritize AGEI training events (identified by the T&E as we described earlier). Because the selection into the study population was based on an individuals' proximity to the threshold score, it was not possible to stratify on AGEI status. However, events that were likely to include more AGEI candidates were purposely oversampled in 2011 and 2012 so as to increase the number of AGEI candidates in the study population.
[12] Eighty percent of EF training events occurred during these four months.
[13] In 2010, because a complete event listing was not available in advance, the events were chosen by convenience, based on scheduling and accessibility.
[14] The survey team was instructed to drop the event from the sample if there were not at least 3 rejected candidates that fell within 20% of the threshold score. In other words, if there were not at least 3 people who could be sampled for the control group, the event was dropped from the sample.

month (equivalent to about 17 USD). This figure may seem low, since it represents the average earnings over the entire study population of 4677 individuals, including those with zero earnings. Only 17 percent of the 2010-2012 pooled sample earned more than 3000 NRs per month, a level deemed to represent "gainful" employment. Fifteen percent of the sample was already engaged in the same trade for which training they applied (denoted as "trade-specific IGA"), indicating that a significant minority of applicants had been looking to upgrade existing skills. Though not older than men in the sample, women were more likely to be married and to have a child, and had lower employment levels and earnings at baseline.

C. Survey response rates, attrition and program take-up

The response rates were quite high for all follow-up surveys (see Table 3).[15,16] We were able to track and successfully interview 88 percent of the baseline survey respondents, yielding a final sample for analysis of 4,101 individuals.[17] In Table 4, we explore the possibility of "*differential attrition*" and show no evidence to support it. Table 4 shows the results of a panel-based regression with attrition as a dependent variable on a set of covariates and the regression results indicate that attrition is not correlated with treatment status.

D. Eligibility for the Training Program

In each course, applicants with scores above the threshold were assigned to training, while applicants whose scores fall below the threshold were not assigned to training workshops. Immediately following the sampling of applicants and before the results of the selection process were announced, a baseline survey was administered. Three factors comprise the eligibility criteria for all EF-sponsored training programs: age (from 16 to 35), education (below SLC,[18] or less than 10 years of formal education), and low self-reported economic status.[19] Only applicants who meet

---

[15] Because the EF-sponsored training courses vary in length from 1 to 3 months, the follow-up survey examines outcomes 9 to 11 months after the end of the training.

[16] The EF itself conducts follow-up with a sample of participants up to 6 months after the training to verify employment and earnings. Hence, the impact evaluation follow up survey occurs 3-5 months after the treatment group's last contact with the program.

[17] The reasons given for loss to follow-up for the 2010 and 2011 cohorts include: inability to track the household (11%), no one in the household during multiple visits (15%), refusal (8%), and respondent migrated for work within Nepal or abroad (8%), respondent migrated after marriage (10%), or other (40%).

[18] The School Leaving Certificate (SLC) is obtained after successfully passing examinations after the 10th grade. To be eligible, EF applicants must have not taken, or not passed, their SLC exams. This criterion has been loosened for some trades starting in 2012. In practice, because the educational status reported on the application is not verified, this criterion was not perfectly adhered to.

[19] An applicant is considered "economically poor" if they report a non-farm per capita household income of less than 3000 Nepali rupees (NRs) per month or, in the case of farming families, less than 6 months of food sufficiency. Since these self-reports are not verified, and applicants know in advance that they must be "poor" in order to be eligible for the program, it is unclear how well this criterion is adhered to.

all three criteria were viable for short-listing.[20,21] Figure 4 displays a sample ranking form used by training providers. In the next section, we detail the scoring and ranking procedure used by providers. The process for ranking candidates and interviewing shortlisted candidates follow streamlines guidelines, including a detailed scoring rubric, instructions for ranking the shortlisted candidates by score, and selecting the top-scoring candidates for participation. The individual score used in ranking candidates consisted of five components: trade-specific education, economic status, social caste, geographic area, and interview score. On each of these components, individuals were scored and each component had a weight assigned to each category. Each individual then had a calculated score by summing across components. Possible scores range from 0 to 100. Although eligibility for training based on the actual score influenced the likelihood of training course enrollment, individual assignment to training was not automatic as it was envisioned because of likely provider manipulation of the component scores, an issue which we tackle below.

III. Empirical Strategy

A. Estimating Treatment Effects of the Training Intervention

In an ideal case, we would be able to examine the effect of training provision on outcomes by using the individual scores assigned by the providers during the interview procedure. The discontinuity in training assignment induced by the threshold score in theory should cause an exogenous change in the probability of training holding individual characteristics constant. However, as mentioned previously, we have reason to assume that training providers were influencing the assigned scores – possibly in response to the payment structure, which rewards completed trainings and trainee placement over drop-outs and non-placed trainees. Hence, manipulating the individual scores is likely to be related to unobserved individual characteristics, and therefore likely to bias the estimates of interest (McCrary 2008).

We follow the approach by Miller et al. (2013), who seek to overcome this challenge by reconstructing the 'actual' individual-specific score from survey data. Currie and Gruber (1996a, b), Cutler and Gruber (1996), and Hoxby (2001) also follow this approach. We then use the reconstructed score to instrument for training enrollment employing fuzzy regression discontinuity

[20] Training providers were advised to shortlist at least 50 percent more candidates than the number of spaces available in the training event.
[21] In addition, and as mentioned earlier, the guidelines also set out a progressive payment structure to incentivize T&E providers to select trainees from particular socially disadvantage groups. The guidelines also allow for T&E providers to select up to 2 "alternates" per event, in case a selected trainee declines to join the program.

set-up similar to the one proposed by Hahn, Todd, and Van der Klaauw (2001). As we cannot rule out that the threshold scores, which were to determine assignment to training, were also not affected by some type of provider manipulation, we re-estimate the threshold scores for each course following the approach proposed by Miller et al. (2013).[22] Specifically, we estimate the following first-stage equation:

$$Trained_i = \alpha + \gamma AboveThreshold_{ic} + \gamma TotalScore_i + \gamma RelativeScore_{ic} + \epsilon_i \text{ , } (1)$$

where $Trained_i$ is an indicator for whether or not an applicant i has received training, $AboveThreshold_{ic}$ is an indicator for the reconstructed assignment score of the applicant being greater or equal to the estimated threshold score of the respective course c he or she applied to, $TotalScore_i$ is the applicant's reconstructed assignment score, and $RelativeScore_{ic}$ is the difference between an applicant's reconstructed assignment score and the estimated threshold score of the course.

We then estimate the following second stage:

$$\Delta Y_i = \varphi + \lambda\, Trained_i + \lambda\, TotalScore_i + \lambda\, RelativeScore_{ic} + \varepsilon_i \text{ , } \qquad (2)$$

where we capture the relationship between the first-differenced outcome Y_i and program enrollment by estimates of λ. To adhere transparently to the identifying assumption that individuals with simulated training eligibility scores very near the threshold are comparable with the exception of their eligibility, we conservatively focus on individuals whose calculated scores lie within two index points of the estimated cutoff (we run the same procedure within three-index, four-index, five-index, and ten-index points and our estimates persist across these various bandwidths).

B. Heterogeneous Effects

[22] The authors follow Chay, McEwan, and Urquiola (2005).

Because treatment heterogeneity has important policy implications, we estimate heterogeneous local average treatment effects (HLATE) based on the framework proposed by Becker et al. (2013). In particular, we estimate a 2SLS equation similar to specification (2):

$$\Delta Y_i = \varphi + \lambda\, Trained_i + \lambda\, H_i x Trained_i + \lambda\, H_i + \lambda\, TotalScore_i + \lambda\, RelativeScore_{ic} + \varepsilon_i\,, \qquad (3)$$

where H_i is an indicator for subgroup. We use the predicted probability of training and its interaction with the subgroup indicator as instruments for *Trained*$_i$ and *H*$_i$*xTrained*$_i$.[23]

IV. Results

A. Balance across Discontinuous Eligibility

Our empirical approach assumes that no individual characteristics, other than vocational training enrollment, that could influence the outcomes of interest vary continuously across our estimated eligibility thresholds. To test this assumption, Table 5 shows results obtained by estimating equations (1) and (2) for individual attributes that could not reasonably change in response to training enrollment (age, ethnicity, gender, or educational attainment among adults). In Figure 5, we present evidence that there is no distinguishable difference around the threshold of the running variable for individual attributes unrelated to treatment. Consistent with our assumption, estimates are not generally distinguishable from zero.

B. Probability of Treatment Assignment and Continuity of Interaction Variables Around the Threshold

To show that probability of treatment jumps at cut-off of the individual training score, we exhibit the probability of treatment assignment in Figure 6. As discussed previously because treatment is not solely determined by the cutoff rule, we see a probability of treatment jump by less than one. We, further, show in Figure 5 and Figure A1 that the subgroup indicators we use to

[23] To predict the probability of training, we estimated a Probit model regressing the training indicator on the subgroup indicator H_i, the assignment indicator $AboveThreshold_{ic}$, an interaction of the two, as well as the total and the relative score variables.

determine heterogeneous treatment effects (applicants gender and trade of training) are continuous across the threshold. This confirms that assignment status is not correlated with interaction variables conditional on the relative assignment score (which we control for in all specifications).

C. Impacts on Employment and Earnings for the Full Sample

In this section, we present the impact on the combined 2010, 2011 and 2012 samples with bandwidth of 2, which minimizes the mean-square of error of the estimator. However, we present a number of robustness checks in section V.

[Table 8 about here]

Table 8 shows the 2SLS results on employment and earnings for the pooled 2010, 2011 and 2012 cohorts. We find no evidence (results in the first row of Table 9) of program impact on the employment rate.[24] Restricting the employment to non-farm activities, we find a significant increase: the rate of participation in non-farm income-generating activities increases by 28 percentage points (from a base of 29.6 percent). Translating the results in percentage change terms, we find that the program increased non-farm employments by 95 percent. These impacts are not only statistically significant but also economically meaningful. We also examine the trade-specific income generating activity (IGA) rate – the percent of individuals who find employment in the same trade as the training that they applied for – and we find impacts of 39 percentage points. The trade-specific IGA impacts are larger than the non-farm employment impacts, suggesting that members of the counterfactual group, even when able to find employment, were less able than the treatment group to find employment in the trade in which they sought training. The EF program also leads to persistent improvements in the underemployment rate (*i.e.,* cases in which people are working fewer hours than they wish). Table 8 shows that EF-sponsored training courses increased hours worked in IGAs for the pooled cohorts by 62 hours per month (*i.e.,* 89 percent).

We detect strong program impacts on monthly earnings. We measure earnings as an individual's total earnings in the past month, including income from all IGAs, but not including unearned income.[25] We observe a statistically significant (at the 1 percent level) increase in

[24] We measure employment by whether the respondent reported any income-generating activities in the past month or not.
[25] If an individual did not work in the past month, his/her earnings are recorded as zero.

monthly earnings for the treatment group by 2,169 NRs (≈ 29 USD), from a baseline average of 1,272 NRs (≈ 17 USD).[26] In percentage terms, this earnings increase translates to a 171 percent for the pooled sample.

With alternative measurements of earnings, we also detect large program impacts. To account for the highly skewed nature of earnings distributions, we examine the impact on logged earnings and we find impacts of approximately 82 percent. A third way to examine the impact on earnings is to consider the proportion of participants who earned a "decent living." The Employment Fund considers 3000 NRs per month (≈ 40 USD) as "gainful employment" and considers this amount as "being productively employed." At baseline, only about 20 percent of the sample was "gainfully employed". The EF training program increased the "gainful employment" rate increases by 31 percentage points.

D. Trade-wise Program Impacts on the Full Sample

The Employment Fund sponsors about 600 training courses annually -- from short four-week courses on incense-stick rolling to three-month technical courses. Table 7 shows the breakdown of courses by trade.

We grouped training courses into seven categories. The most common categories of training in our sample are Electrical/Electronics/Computer (*e.g.,* electric wiring, computer hardware technician, and mobile phone repair), Construction/Mechanical/Automobile (*e.g.,* arc welding, brick molding, furniture making, motor bike service), and Tailoring/Garment/Textile (*e.g., galaicha* weaving, garment fabrication, hand embroidery, tailoring and dressmaking). Because we sample approximately the same number of applicants per event, the breakdown of applicants by trade (Panel 2 of Table 7) is very similar to the event-wise breakdown.

Table 9 shows that the impacts of the skills training program differed markedly by type of trade for the pool 2010-2012 samples. Training in electronics, beautician, and tailoring consistently show strong impacts on employment—graduates of these training programs are more likely to have employment in general and are also more likely to be working in the trade in which they were trained. Beautician training shows large impacts on both employment and earnings.

[26] This average is based on the entire study cohort, including those with zero earnings at baseline. The average earnings among those with non-zero earnings were 2928 NRs, translating to a percentage increase in earnings of 30%.

[Table 9 about here]

We detect impact on earnings for handicraft training and no significant impacts on employment or earnings outcomes for the remaining three trades. Results for construction show no significant impacts. Overall, the results in Tables 9 reveal substantial heterogeneity in employment outcomes across the various types of training. The positive and significant impacts discussed previously are driven almost entirely by three categories of trades: electronics, beautician training, and tailoring trades show positive and significant impacts on employment and earnings across the three cohorts. We find no impacts for the food/hospitality and construction training. Handicraft-related trainings showed positive and significant earnings impacts, but the effects are not consistent across all outcomes.

E. Gender- and Age- disaggregated Impacts

We also explore program impacts for men and women (shown in Table 10). To that end, Tables 10 disaggregate the results to compare outcomes for men versus women. The employment impacts are significantly larger for women than for men. The results for other economic outcomes, such as hours worked, earnings, and type of employment, are almost double the magnitude for females and the only program impact that is statistically significant for males are on trade-specific labor force participation, whereas for females all program impacts are statistically significant.

[Table 10 about here]

Several factors could account for the differential impacts in employment outcomes by gender. First, when asked, training providers suggested that female students attend classes more and are more diligent than male students. We lack data on the attendance or completion rates of EF trainees; however, the rate of non-compliance with treatment assignment is equal for men and women, suggesting that this mechanism is unlikely to exert much influence on outcomes. Second, the Employment Fund introduced life skills training for women in 2011 in all of its training courses.[27] Because all women received life-skills training, we cannot disentangle the influence of

[27] The forty-hour curriculum covered topics such as negotiation skills, workers' rights, sexual and reproductive health, and dealing with discrimination. Female students overwhelmingly responded positively to the life skills training, often claiming that it was one of their favorite parts of the course. The skills learned and the positive experience in this life skills training may contribute to the increased employment impact for women, which is line with the advice from experts in vocational training from around the world, who increasingly advocate for the inclusion of life skills in technical training programs

this factor, from other program elements, on outcomes. A third explanation could relate to men start with a higher level of non-farm employment (47 percent compared to 20 percent for women at baseline) and therefore it may be easier for women to make large gains on the extensive margin. A fourth possible explanation relates to the difference between the types of trades that men and women apply for. Although the Employment Fund specifically tried to encourage female participation in non-traditionally female trades, most of the training courses tend to be heavily gender-segregated. For example, men tend to dominate electronics and construction courses, while the tailoring and beautician trainings are comprised almost exclusively of females. As shown earlier, the tailoring and beautician trainings exhibit the largest employment impacts.

V. Robustness Checks

A. Bandwidth Selection

To investigate the robustness of our results relative to bandwidth choice, we estimate a variety of alternative non-parametric specifications based on our main estimating equations. We re-estimate equations (3) and (4) bandwidth within 3 index scores of threshold. Appendix Tables A1-A3 present the results and show that all specifications are stable both in statistical significance and coefficient magnitude. In Table 11 we present estimates from alternative bandwidth choices -- 2, 3, 4, 5 and 10 index scores of threshold. Table 11 shows stability in coefficient magnitudes and levels of statistical significance.

[Table 11 about here]

B. Combined Difference-in-Differences and Propensity Score Approach

As a robustness check for estimates of program impact, we also employ a combination of difference-in-difference with a propensity score matching technique (Campbell 1969; Meyer 1995). This approach in the context of training programs has the potential to purge potential differences between observable characteristics for trainees and non-trainees following Dehejia and Wahba (2002). We estimate:

$$Y_{it} = \alpha + \delta\,(A_{it}\,x\,Treat_{it}) + u_i + v_t + \varepsilon_{it} \tag{3}$$

Y_{it} is the employment outcome of interest for individual i from training event j at time t; $Treat_i$ is an indicator, equal to 1 for the treatment group and 0 for control; A_{it} is an indicator for the period when treatment occurs; u_i captures program effects; v_t captures the time effects, ε_{it}, is an idiosyncratic error term, clustered by training event. After estimating a propensity score[28], we derive the estimated treatment effect using two methods: "inverse propensity score weighting" (IPSW) and nearest neighbor matching (NN). In the IPSW method individuals are weighted according to the inverse of their estimated propensity to participate in the program. The weighted observations are then used in a DID regression, as given by equation (3). We present this method estimates under the IPSW specification in the Appendix B tables.[29,30] The NN matching algorithm, in which each individual in the treatment group is compared to a fixed number of control observations (in our estimation we use four observations) with the closest propensity score. We present "NN specification" in Appendix B. Following Smith and Todd (2005), we estimate the difference-in-difference matching estimator for the training program effect δ as follows:

$$\widehat{\delta_M} = \frac{1}{N_T}\sum_{i\epsilon T}\left[\left(y_{it_1} - y_{it_0}\right) - \sum_{j\in C} W_{ij}\left(y_{jt_1} - y_{jt_0}\right)\right] \tag{4}$$

N_T is the number of treatment observations, the subscript $t_1$denotes follow-up observations and t_0 denotes baseline observations; W_i is a matrix of weights. Weights for nearest-neighbor matching are computed by:

$$W_{ij}\left(y_{jt_1} - y_{jt_0}\right) = \frac{1}{x}\sum_{\substack{j=1\\ j\in A_x}}^{x}\left(y_{jt_1} - y_{jt_0}\right) \tag{5}$$

[28] We employ various specifications, including the individual training score, the individual training score and demographic variables, the five subinterview scores and finally the demographic variables plus provider/district/cohort/city fixed effects. The results are stable across various specifications though we report the last specification based on demographic variables plus provider/district/cohort/city fixed effects because that specification yields the best overlap of treatment and control unit distributions in the common support area.

[29] We implement IPSW following Hirano et al (2003).

[30] This particular weighting method, as opposed to matching approaches, has the nice property of including all the data (unless weights are set to 0) and does not depend on random sampling, thus providing for replicability. We use a weighted least squares regression model, with weights of $1/\hat{\pi}$ for the treatment group and $1/(1-\hat{\pi})$ for the control group, where $\hat{\pi}$ is the estimated propensity score from (2). Standard errors are clustered by training event.

A_x is a set of x observations with the lowest values of $|\hat{\pi}_i - \hat{\pi}_j|$. As in the two previous models outlined in this section, the dependent variable is the first difference of a given outcome between the baseline observation and the follow-up observation. We measure outcomes approximately one year after the start of training.[31, 32]

Appendix Table B11 shows the ATT results on employment and earnings for the pooled 2010, 2011 and 2012 cohorts based on the combined difference-in-difference and propensity score matching techniques. Unlike the RDD results, in this specification, we detect strong evidence of consistent impact on the employment rate across all specifications.[33] All three models indicate a positive and significant effect, despite the high employment rate (*i.e.,* 61 percent) at baseline. Restricting the employment to non-farm activities, we also find a significant increase: the rate of participation in non-farm income-generating activities increases by 21 percentage points (from a base of 29.6 percent). Translating the results in percentage change terms, we find that the program increased non-farm employments by 71 percent. These impacts are not only statistically significant but also economically meaningful. We detect strong program impacts, though smaller in impacts than revealed by the RDD approach, on monthly earnings. We observe a statistically significant (at the 1 percent level) increase in monthly earnings for the treatment group by 976 to 1099 NRs (≈ 14 USD), from a baseline average of 1272 NRs (≈ 17 USD).[34] In percentage terms, this earnings increase translates to a 81 percent for the pooled sample. The impact on logged earnings is a little over 100 percent. The EF training program increased the "gainful employment" rate (i.e., the rate of new employment with earnings over 3000) increases by 16 to 17 percentage points, a result statistically significant across all three models. We also examine the trade-specific income generating activity (IGA) rate – the percent of individuals who find employment in the same trade

[31] Because the EF-sponsored training courses vary in length from 1 to 3 months, the follow-up survey examines outcomes 9 to 11 months after the end of the training.
[32] First, we address concerns about pre-existing differences and time-varying trends that could account for observed training effects when comparing trainees and non-trainees.[32] Table B1 presents baseline participant characteristics (*i.e.*, balancing tests) for a set of 38 demographic indicators. These tests are based on "ITT" comparisons of the treatment group (*i.e.*, individuals whose scores qualify them for admission to an EF training event) and the control group. The baseline balance tests for the pooled sample (2010-2012) indicate that significant differences exist between treatment and control groups for baseline observable characteristics and pre-treatment outcome variables.[32] Relative to rejected candidates, treated individuals are more likely to be Janajati and are less likely to have finished SLC (10th grade), characteristics which reflect the eligibility criteria and the EF"s differential pricing scheme for vulnerable groups. Further, the likelihood of treated individuals being engaged in non-farm and trade specific employment before take up of training was higher, as well as their working hours and ability to earn more than 3000 NRs a month. These differences are consistent with training providers' incentives to select candidates they think will perform best. Finally, individuals in the treatment group are also less likely to have control over savings and money of their own at baseline. To address these differences (and potential differences in unobservable characteristics) we applied a difference-in-difference approach in our analysis. However, growth in outcome variables and the may not follow a common trend, particularly when starting off at very different initial levels. Although it does not resolve the parallel trend assumption, we additionally applied propensity score weighting and matching techniques to achieve a higher degree of baseline comparability across groups.
[33] We measure employment by whether the respondent reported any income-generating activities in the past month or not.
[34] This average is based on the entire study cohort, including those with zero earnings at baseline. The average earnings among those with non-zero earnings were 2928 NRs, translating to a percentage increase in earnings of 30%.

as the training that they applied for – and we find impacts of 24 percentage points. The trade-specific IGA impacts are larger than the non-farm employment impacts, suggesting that members of the control group, even when able to find employment, were less able than the treatment group to find employment in the trade in which they sought training. Based on the propensity score approach, we find that the EF program leads to persistent improvements in the underemployment rate (*i.e.,* cases in which people are working fewer hours than they wish). Table B11 shows that EF-sponsored training courses increased hours worked in IGAs for the pooled cohorts by 30-31 hours per month (*i.e.,* 44 percent). All three model specifications exhibit a statistically significant and positive impact.

Table B12 shows results for program impact heterogeneity. The impacts of the skills training program differed markedly by type of trade for the pool 2010-2012 samples. Consistent and exactly aligned with the results based on the RDD approach, training in electronics, beautician, and tailoring consistently show strong ATT impacts on employment—graduates of these training programs are more likely to have employment in general and are also more likely to be working in the trade in which they were trained. Beautician training shows large impacts on both employment and earnings. We detect no significant impacts on employment or earnings outcomes for the remaining four trades. Results for food and hospitality (*e.g.,* cooking and wait service) show no significant ATT impacts; results for construction show no significant impacts except for a marginal impact on trade-specific employment and on earning more than 3000 NRs per month. For the remaining three trades (*i.e.,* poultry technician, handicrafts and farming), we detect some ATT impacts but they are not consistent across models. Overall, the results in Tables B12 reveal substantial heterogeneity in employment outcomes across the various types of training. The positive and significant impacts are driven almost entirely by three categories of trades: electronics, beautician training, and tailoring trades show positive and significant impacts on employment and earnings across both cohorts. We find no impacts for the food/hospitality and farming training. Construction-related trainings showed positive and significant impacts, but the effects are not consistent across outcomes.

Finally, we show program impacts for men and women (shown in Table B13). To that end, Tables B13-B14 disaggregate the results to compare outcomes for men versus women, and for younger women (the “AGEI” population) versus older women. Corroborating the RDD findings, Tables B13-B14 show the employment impacts are larger, almost double the magnitude for women than they are for men. The results for other economic outcomes, such as hours worked, earnings, and type of employment, are similar for both sexes.

VI. Conclusion

Training interventions have been hailed as one potential solution to facilitate youth's transition to productive employment and higher earnings. Using a regression discontinuity method in the context of a large vocational training program in Nepal, we find positive and statistically significant effects on labor market outcomes for training program participants on employment rates (any or non-farm), finding employment related to the skill they learned, hours worked, earnings, and the proportion of people earnings more than 3000 NRs per month.

Individuals selected for EF training programs experience an increase in non-farm employment of 28 percentage points for an overall gain of 93 percent. We detect an increase in average monthly earnings of approximately 71 percent. Alongside the sizable general impacts on employment outcomes, we find that training courses in electronics, beautician services, and tailoring underpin most of the EF program's impacts. These three categories of training are much more effective in consistently increasing employment and earnings than construction, poultry rearing, handicrafts, and food preparation and hospitality.

Perhaps most strikingly, we find larger impacts on employment for women than for men. Women selected for training in 2010 to 2012 experience overall and non-farm employment gains of 20 and 40 percentage points respectively, while the corresponding impacts for men are 12 and 10 percent. We find significant differences by gender also on other economic indicators such as trade-specific employment, logged earnings, and gainful employment.

Our estimates of the employment effects of this training intervention are among the largest for training programs around the world. Although pinpointing the exact mechanisms is an important topic for future studies, we posit two potential explanations: first, the EF had time to become established and to develop systems prior to the launch of the training intervention in 2010. In this intervention, the program was already operating at scale, and the service delivery processes already road-tested; second, the training program was designed around employment outcomes in that training providers had to complete market assessments to ensure future employability in the trades in which they proposed to train individuals. While these are important caveats to the conclusion that Nepal's training program these generates significant employment improvements, our results suggest an intervention model of that may have general applications for designing effective labor market interventions elsewhere.

References

Adoho, Franck, Shubha Chakravarty, Dala T. Korkoyah Jr., Mattias Lundberg, and Afia Tasneem. 2013. "The Impact of an Adolescent Girls Program: The EPAG Project in Liberia" , Working Paper.

Aedo, Cristián and Sergio Nuñez. 2004. "The Impact of Training Policies in Latin America and the Caribbean: The Case of Programa Joven" , Inter-American Development Bank Working Paper #R-483.

Almeida, Rita, Jere Behrman, and David Robalino. 2012. *The Right Skills for the Job?: Rethinking Training Policies for Workers*: World Bank Publications.

Alzúa, María L., Guillermo Cruces, and C. Lopez Erazo. 2013. "Youth Training Programs Beyond Employment. Evidence from a Randomized Controlled Trial" *La Plata, Buenos Aires: CEDLAS.(Mimeo.)*.

Anderson, M. L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects" *Journal of the American Statistical Association*, 103(484).

Angelucci, Manuela and Orazio P. Attanasio. 2006. "Estimating ATT Effects with Non-Experimental Data and Low Compliance." , No. 2368. IZA Discussion Papers.

Attanasio, Orazio, Adriana Kugler, and Costas Meghir. 2011. "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial" *American Economic Journal: Applied Economics*, 3(3): 188-220.

Attanasio, Orasio, Adriana Kugler, and Costas Meghir. 2008. "Training Disadvantaged Youth in Latin America: Evidence from a Randomized Trial" , NBER Working Paper 13931.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2012. "Empowering Adolescent Girls: Evidence from a Randomized Control Trial in Uganda" *London: London School of Economics*.

Betcherman, Gordon, Amit Dar, and Karina Olivas. 2004. *Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries*: Social Protection, World Bank.

Betcherman, Gordon, Martin Godfrey, Olga Susana Puerto, Friederike Rother, and Antoneta Stavreska. 2007. "Global Inventory of Interventions to Support Young Workers: Synthesis Report" *World Bank*, Global Inventory of Interventions to Support Young Workers: Synthesis Report.

Blattman, Christopher and Jeannie Annan. 2016. "Can Employment Reduce Lawlessness and Rebellion? A Field Experiment with High-Risk Men in a Fragile State" *American Political Science Review*, 110(1): 1-17.

Bryson, A., R. Dorsett, and S. Purdon. 2002. "The use of Propensity Score Matching in the Evaluation of Labour Market Policies" , Working Paper No. 4, Department for Work and Pensions.

Burghardt, John and Peter Z. Schochet. 2001. "National Job Corps Study: Impacts by Center Characteristics." .

Campbell, Donald T. 1969. "Reforms as Experiments" *American Psychologist*, 24: 407–429.

Card, David and Daniel M. Sullivan. 1987. *Measuring the Effect of Subsidized Training Programs on Movements In andOut of Employment*.

Card, D., J. Kluve, and A. Weber. 2009. "Active Labour Market Policy Evaluations: A Metaanalysis" , Austrian Center for Labour Economics and the Analysis of the Welfare State Working Paper, (0902).

Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas, and Yuri Soares. 2011. "The Labor Market Impacts of Youth Training in the Dominican Republic." *Journal of Labor Economics*, 29(2): 267-300.

Carneiro, Pedro M. and James J. Heckman. 2003. "Human Capital Policy" .

Chong, Alberto and Jose Galdo. 2006. "Does the Quality of Training Programs Matter? Evidence from Bidding Processes Data" .

Currie, Janet and Jonathan Gruber. 1996. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health" *The Quarterly Journal of Economics*, 111(2): 431-466.

------. 1996. "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women" *Journal of political Economy*, 104(6): 1263-1296.

Cutler, David M. and Jonathan Gruber. 1996. "Does Public Insurance Crowd Out Private Insurance?" *The Quarterly Journal of Economics*, 111(2): 391-430.

Dehejia, Rajeev H. and Sadek Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies" *Review of Economics and statistics*, 84(1): 151-161.

Elias, Victor, Fernanda Ruiz-Nunez, Ricardo Cossa, and David Bravo. 2004. "An Econometric Cost-Benefit Analysis of Argentina's Youth Training Program" *Inter-American Development Bank, Research Network Working Paper*.

Fella, G. and G. Gallipoli. 2007. "Education and Crime Over the Life Cycle" , Rimini Center 15-07.

Frese, Michael and Gisela Mohr. 1987. "Prolonged Unemployment and Depression in Older Workers: A Longitudinal Study of Intervening Variables" *Social science & medicine*, 25(2): 173-178.

Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Vermeersch. Christel M.J. 2011. *Impact Evaluation in Practice*: World Bank.

Glazerman, Steven, Dan M. Levy, and David Myers. 2003. "Nonexperimental Versus Experimental Estimates of Earnings Impacts" *The Annals of the American Academy of Political and Social Science*, 589(1): 63-93.

Goldsmith, Arthur H., Jonathan R. Veum, and William Darity. 1997. "The Impact of Psychological and Human Capital on Wages" *Economic inquiry*, 35(4): 815-829.

González-Velosa, Carolina, Laura Ripani, and David R. Shady. 2012. "¿ Cómo Mejorar Las Oportunidades De Inserción Laboral De Los Jóvenes En América Latina?" *Technical Note IDB-TN305, Labor Market and Social Security Unit, Inter-American Development Bank, Washington DC*.

Groh, Matthew, Nandini Krishnan, David J. McKenzie, and Tara Vishwanath. 2012. "Soft Skills Or Hard Cash? the Impact of Training and Wage Subsidy Programs on Female Youth Employment in Jordan" *The Impact of Training and Wage Subsidy Programs on Female Youth Employment in Jordan (July 1, 2012).World Bank Policy Research Working Paper*(6141).

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a regression-discontinuity Design" *Econometrica*, 69(1): 201-209.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1998. "Matching as an Econometric Evaluation Estimator" *The Review of Economic Studies*, 65(2): 261-294.

------. 1998. "Matching as an Econometric Evaluation Estimator" *The Review of Economic Studies*, 65(2): 261-294.

Heckman, James J., Alan B. Krueger, and BM Friedman. 2004. "Inequality in America" *Cambridge (Mass.)*.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs" *Handbook of labor economics*, 3: 1865-2097.

Heckman, J., R. LaLonde, and Smith J. 1999. "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics* , ed. O. Ashenfelter and David Card, 1865-2097. Vol. III. Elsevier, Amsterdam.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score" *Econometrica*, 71(4): 1161-1189.

Holzer, Harry J. 2007. "What might Improve the Employment and Advancement Prospects of the Poor?" Paper presented at Federal Reserve Bank of Chicago Conference on Strategies for Improving Economic Mobility of Workers, Chicago, November.

Hopenhayn, Martin. 2002. "Youth and Employment in Latin America and the Caribbean: Problems, Prospects and Options" , ECLAC Working Paper.

Hoxby, Caroline M. 2001. "All School Finance Equalizations are Not Created Equal" *The Quarterly Journal of Economics*, 116(4): 1189-1231.

Youth Employment Statistics in Nepal. 2014. . http://apyouthnet.ilo.org/stats/youth-employment-statistics-in-nepal/at_download/file1 (accessed September/1).

Kling, Jeffrey R. and Jeffrey B. Liebman. 2004. *Experimental Analysis of Neighborhood Effects on Youth*: John F. Kennedy School of Government, Harvard University.

Kluve, Jochen. 2006. "The Effectiveness of European Active Labor Market Policy" , IZA Discussion Paper No. 2018; RWI Discussion Paper No. 37.

LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data" *The American Economic Review*: 604-620.

------. 1995. "The Promise of Public Sector-Sponsored Training Programs" *The Journal of Economic Perspectives*, 9(2): 149-168.

Leuven, E. and B. Sianesi. 2003. "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing and Covariate Imbalance Testing" , 4.0.6 17may2012.(http://ideas.repec.org/c/boc/bocode/s432001.html).

Linn, M. W., R. Sandifer, and S. Stein. 1985. "Effects of Unemployment on Mental and Physical Health" *American Journal of Public Health*, 75(5): 502-506.

Macours, Karen. 2011. "Increasing Inequality and Civil Conflict in Nepal" *Oxford Economic Papers*, 63(1): 1-26.

Maitra, Pushkar and Subha Mani. 2017. "Learning and Earning: Evidence from a Randomized Evaluation in India" *Labour Economics*, 45: 116-130.

------. 2017. "Learning and Earning: Evidence from a Randomized Evaluation in India" *Labour Economics*, 45: 116-130.

Maitra, Pushkar and Subha Mani. 2014. "Learning and Earning: Evidence from a Randomized Evaluation in India." *GCC Working Paper Series, GCC 14-05*.

Malamud, Ofer and Cristian Pop-Eleches. 2010. "General Education Versus Vocational Training: Evidence from an Economy in Transition" *The review of economics and statistics*, 92(1): 43-60.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test" *Journal of Econometrics*, 142(2): 698-714.

Meyer, Bruce D. 1995. "Natural and Quasi-Experiments in Economics" *Journal of Business and Economic Statistics*, 13(2): 151–161.

Ñopo, Hugo, Jaime Saavedra-Chanduví, and Miguel Robles. 2007. "Occupational Training to Reduce Gender Segregation: The Impacts of ProJoven" , Inter-American Development Bank Working Paper #623.

O'Brien, Peter C. 1984. "Procedures for Comparing Samples with Multiple Endpoints" *Biometrics*: 1079-1087.

Reis, Mauricio. 2015. "Vocational Training and Labor Market Outcomes in Brazil" *The BE Journal of Economic Analysis & Policy*, 15(1): 377-405.

Robalino, D., Friederike Rother, David Newhouse, and Mattias Lundberg. 2013. "Youth Unemployment: A Development Agenda for the Next Decade" , World Bank Social Protection and Labor Discussion Paper no.1308.

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

------. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects" *Biometrika*, 70(1): 41-55.

Rubin DB. 1980. "Randomization Analysis of Experimental Data in the Fisher Randomization Test." *Journal of the American Statistical Association*, 75: 591–593.

Rubin, Donald B. and Neal Thomas. 1996. "Matching using Estimated Propensity Scores: Relating Theory to Practice" *Biometrics*: 249-264.

------. 1996. "Matching using Estimated Propensity Scores: Relating Theory to Practice" *Biometrics*: 249-264.

Smith, Jeffrey and Petra E. Todd. 2005. "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators" *Journal of Econometrics*(125): 305–353.

Solomon, Goosain, Michael Frese, Christian Friedrich, and Matthias Glaub. 2013. "Can Personal Initiative Training Improve Small Business Success? A Longitudinal South African Evaluation Study" *The International Journal of Entrepreneurship and Innovation*, 14(4): 255-268.

Empowering Girls and Young Women in Liberia. 2014. The World Bank. http://siteresources.worldbank.org/EXTGENDER/Resources/AGILiberiaBrief.pdf (accessed September/1).

International Labour Organization, ILOSTAT Database. 2017. The World Bank. http://data.worldbank.org/indicator/SL.UEM.1524.ZS (accessed July 10).

New Work Opportunities for Women in Jordan. 2014. The World Bank. http://siteresources.worldbank.org/EXTGENDER/Resources/AGIJordanBrief.pdf (accessed September/1).

Todd, Petra E. 2013. "How to Improve Women's Employability and Quality of Work in Developing and Transition Economies" *http://www.womeneconroadmap.org/sites/default/files/Todd_Improve%20Women%20Employability%20and%20Quality%20of%20Work.pdf*.

World Population Monitoring. 2017. United Nations. http://www.un.org/en/development/desa/population/publications/pdf/fertility/12_66976_adolescents_and_youth.pdf (accessed July 10).

World Bank. 2013. "World Development Report 2013: Jobs." *World Bank Publications*.

FIGURE 1. PROJECT TIMELINE AND SAMPLE SIZE

	2010			2011			2012		
2010 Cohort (N=1556; 1184 Treatment, 372 Control)	Baseline Survey	Training		First Follow-up Survey					
2011 Cohort (N=1586; 1237 Treatment, 349 Control)				Baseline Survey	Training		First Follow- up Survey		
2012 Cohort (N=1535; 1044 Treatment, 491 Control)							Baseline Survey	Training	

FIGURE 2. DISTRICTS COVERED IN 2010-2011

FIGURE 3. DISTRICTS COVERED IN 2012

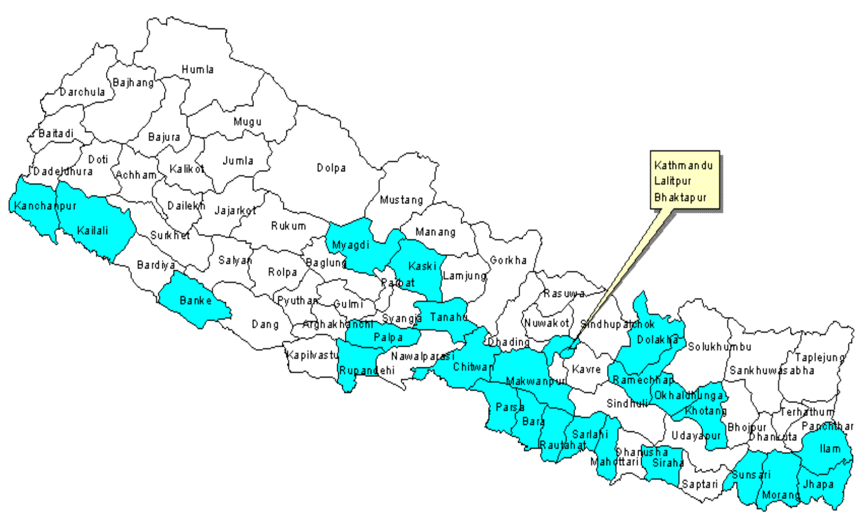


FIGURE 4: EXAMPLE RANKING FORM

#	Name and Surname	Immediate contact telephone	Entry Requirement (Y/N)			Selection Criteria (Individual Scores) 1-4. Short-listing (70%)				5. Interview (30%)	Final Marks	Rank
			Age 16-35 *(Write age)*	Education < SLC	<6 mon. food sufficiency / < Rs. 3,000 mthly income	1. Trade-specific education (15)	2. Economic status (20)	3. Social caste (25)	4. Geographical rep (10)		**TOTAL** MRAKS (100) *List candidates from highest to lowest*	
1	Jane Doe 1	12345678	21	Y	Y	15	20	20	5	26	86	1
2	John Doe 1	12345678	35	Y	Y	15	20	20	5	26	86	2
3	Jane Doe 2	12345678	23	Y	Y	15	20	20	5	25	85	3
4	John Doe 2	12345678	16	Y	Y	15	20	20	5	25	85	4
5	Jane Doe 3	12345678	27	Y	Y	15	20	20	5	23	83	5
6	John Doe 3	12345678	19	Y	Y	15	15	20	5	25	80	6
7	Jane Doe 4	12345678	37	Y	Y	15	15	20	5	25	80	7
8	John Doe 4	12345678	35	Y	Y	15	15	20	5	23	78	8
9	Jane Doe 5	12345678	22	Y	Y	15	15	20	5	23	78	9
10	John Doe 5	12345678	23	Y	Y	15	15	20	5	23	78	10
11	Jane Doe 6	12345678	25	Y	Y	15	15	20	5	23	78	11
12	John Doe 6	12345678	18	Y	Y	15	15	20	5	23	78	12
13	Jane Doe 7	12345678	20	Y	Y	15	15	20	5	23	78	13
14	John Doe 7	12345678	16	Y	Y	15	15	20	5	22	77	14
15	Jane Doe 8	12345678	18	Y	Y	15	15	20	5	22	77	15
16	John Doe 8	12345678	24	Y	Y	15	15	20	5	21	76	16
17	Jane Doe 9	12345678	25	Y	Y	15	15	20	5	21	76	17
18	John Doe 9	12345678	32	Y	Y	15	15	20	5	21	76	18
19	Jane Doe 10	12345678	20	Y	Y	15	15	20	5	18	73	19
20	John Doe 10	12345678	30	Y	Y	15	15	20	5	8	63	20

Note: Red line indicates cut-off between accepted and rejected candidates

FIGURE 5: CONDITIONAL EXPECTATION AT BASELINE FOR COVARIATES

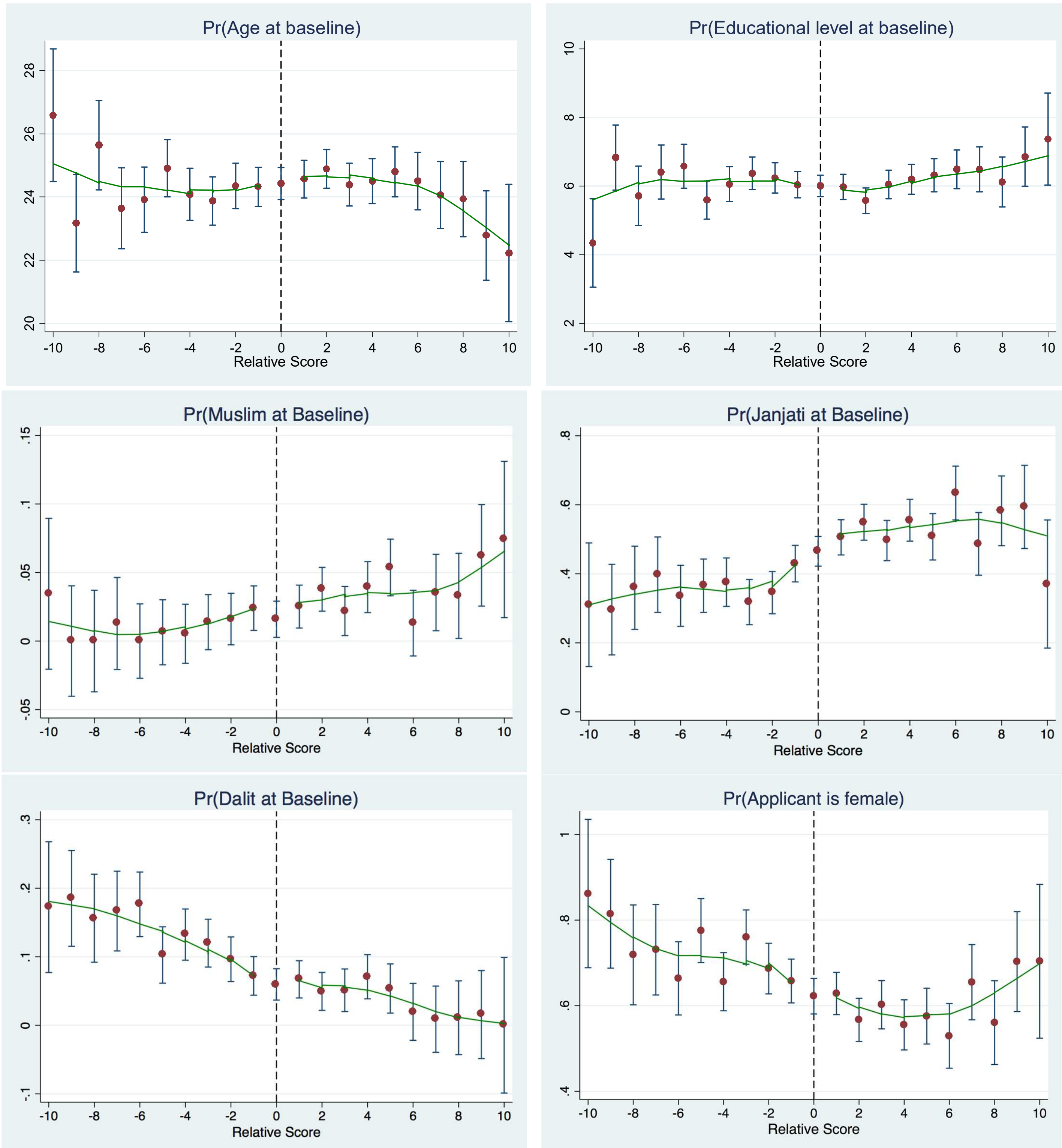


FIGURE 6: PROBABILITY OF TRAINING AT BASELINE

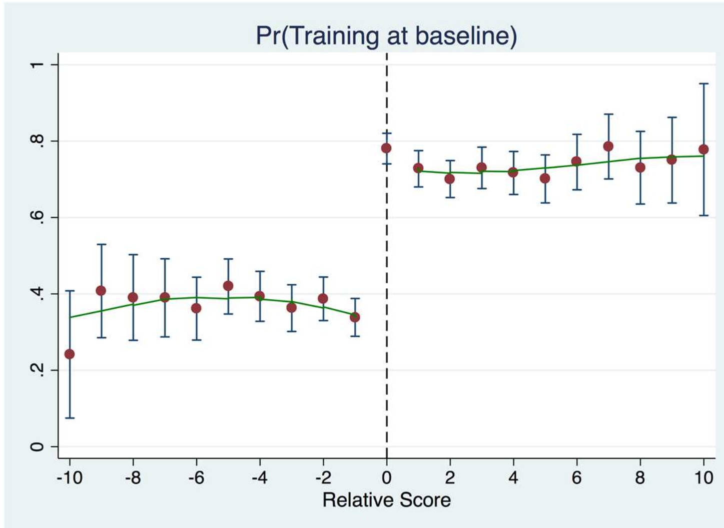


TABLE 1. SCALE EMPLOYEMNT FUND PROGRAM AND AGEI SUB-GROUP

		2010	2011	2012
All EF Programs	Total T&E providers	21	32	35
	Total Events	598	645	711
	Total trained	11750	12869	14255
AGEI Only	T&E providers working with AGEI	11	13	13
	Total Events	110	218	246
	Total trained	808	1664	1936

Notes: T&E is an acronym for "training and employment" providers; AGEI group is women ages 16-24

TABLE 2. SAMPLE SUMMARY OF EVENTS, BASELINE SURVEYS

	2010	2011	2012
Total # events conducted by EF in Jan-Apr	110	142	143
# events randomly sampled	N/A	182	112
# events included in baseline survey	65	69	85
# districts covered	30	34	29
# T&E providers covered	18	26	28

Notes: More events were sampled than conducted in Jan-Apr 2011 because some events that were scheduled for Jan-Apr were delayed and did not start on time.

TABLE 3. SURVEY RESPONSE RATES

	Baseline	**Follow-up**	**Follow-up rate**
2010 cohort			
Above Threshold	1184	1047	88.43%
Below Threshold	372	330	88.71%
Total	1556	1377	88.50%
2011 cohort			
Above Threshold	1237	1113	89.98%
Below Threshold	349	306	87.68%
Total	1586	1419	89.40%
2012 cohort			
Above Threshold	1044	889	85.15%
Below Threshold	491	417	84.93%
Total	1535	1306	89.40%

TABLE 4. CORRELATES SURVEY ATTRITION (2010-2012 POOLED COHORTS)

	(1)	(2)	(3)
"Above Threshold"	0.052	0.042	0.090
	(0.056)	(0.056)	(0.090)
Female		0.472***	0.485***
		(0.116)	(0.168)
AGEI		0.065	0.018
		(0.105)	(0.109)
Parent		0.062	0.006
		(0.104)	(0.107)
Married		0.136	0.084
		(0.101)	(0.101)
Dalit		0.259***	-0.250**
		(0.097)	(0.112)
Janjati		-0.048	-0.039
		(0.062)	(0.066)
Any IGA at baseline		0.204***	0.150**
		(0.061)	(0.062)
Age: under 25		-0.217**	-0.242***
		(0.088)	(0.093)
N	4677	4487	4487
District, TE dummies	No	No	Yes

Notes: All regressions use probit models. "District, TE dummies" indicates that the regression controls for district and training provider effects. Columns 2 and 3 also include training category dummies (not shown). All standard errors are clustered by event.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 5. BALANCE TESTS ON DEMOGRAPHIC CHARACTERISTICS

	(1) Age	(2) Dalit	(3) Janjati	(4) Muslim	(5) Education	(6) Male
2SLS Estimate	-0.486	0.001	-0.019	-0.035	0.292	-0.001
	(1.026)	(0.046)	(0.091)	(0.023)	(0.654)	(0.089)
Observations	1,778	1,778	1,778	1,778	1,778	1,778

Notes: Bandwidth is within 2 index scores of threshold; Standard errors (reported in brackets) clustered at the event level where possible.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 6. TYPE OF TRAINING, 2010-2012 COHORTS

Panel 1: EVENT-WISE TABULATION	**2010**		**2011**		**2012**	
	Number	**%**	**Number**	**%**	**Number**	**%**
Farming	0	0	0	0	5	6
Poultry	2	3	0	0	0	0
Food Prep/ Hospitality	11	17	3	4	2	2
Electrical/ Electronics/Computer	9	14	14	20	14	16
Handicraft & Incense	3	4	4	6	5	6
Construction/Mechanical/Automobile	20	31	13	19	30	35
Beautician /Barber	2	3	5	7	4	5
Tailoring/ Garment/Textile	18	28	30	44	24	28
TOTAL	65	100	69	100	85	99
Panel 2: APPLICANT-WISE TABULATION	**2010**		**2011**		**2012**	
	Number	**%**	**Number**	**%**	**Number**	**%**
Farming	0	0	0	0	92	7
Poultry	41	3	0	0	0	0
Food Prep/ Hospitality	195	14	38	3	32	2
Electrical/ Electronics/Computer	178	13	277	19	186	14
Handicraft & Incense	87	6	79	6	69	5
Construction/Mechanical/Automobile	413	30	258	18	457	35
Beautician /Barber	61	4	117	8	61	5
Tailoring/ Garment/Textile	415	30	650	46	396	30
TOTAL	1390	100	1419	100	1306	99

Notes: This table only includes panel observations (those who were interviewed at baseline and midline).

TABLE 8. EMPLOYMENT, 2010-2012 POOLED COHORTS, BANDWIDTH WITHIN 2 INDEX SCORES

Outcome	Any IGA (1=Yes)	Any non-farm IGA (1=Yes)	Trade-specific IGA (1=Yes)	Hours worked in past month	Earnings	Logged earnings	Earnings > 3000 NRs. (1=Yes)
2SLS Estimate	0.161	0.278**	0.387***	61.90***	2,169***	2.700***	0.308***
	(0.105)	(0.110)	(0.103)	(21.09)	(739.9)	(1.025)	(0.107)
First-stage F-statistic	145.7	145.7	145.7	145.7	138.4	138.4	145.7
Baseline mean	0.612	0.296	0.18	69.261	1271.542	3.291	0.19
	(0.487)	(0.457)	(0.384)	(87.273)	(2197.669)	(3.817)	(0.393)
Observations	1777	1777	1777	1777	1777	1777	1777

Notes: Bandwidth is within 2 index scores of threshold; Standard errors (reported in brackets) clustered at the event level where possible.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 9. EMPLOYMENT BY TRADE, 2010-2012 COHORTS, BANDWIDTH WITHIN 2 INDEX SCORES

	2SLS Estimate			
	any nonfarm IGA	trade-specific IGA	monthly earnings (NRs)	earnings > 3000
	(1)	(2)	(3)	(4)
Full Sample (pooled across all training types)				
Training: Food prep/Hospitality (N=265)	0.469	0.359	2572	0.631
	(0.635)	(0.619)	(4317)	(0.642)
Training: Electrician & Electronics (N=641)	0.278*	0.541***	2617**	0.154**
	(0.172)	(0.168)	(1166)	(0.175)
Training: Handicraft & Incense stick making (N=235)	0.307	0.187	5739**	0.420**
	(0.361)	(0.353)	(2529)	(0.366)
Training: Construction (N=1128)	-0.108	0.109	475.9	0.183
	(0.154)	(0.150)	(1061)	(0.157)
Training: Beautician/Barber (N=239)	0.421*	0.874***	4763***	0.665**
	(0.258)	(0.252)	(1723)	(0.262)
Training: Weaving/Tailoring/Garment Making (N=1461)	0.281*	0.538***	2307**	0.444***
	(0.148)	(0.144)	(1030)	(0.151)
Clustered standard errors (by event)	Yes	Yes	Yes	Yes

Notes: No poultry technician trainings were included in the 2011 sample.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 10. EMPLOYMENT OUTCOMES, BY GENDER, 2010-2012 POOLED COHORTS, BANDWIDTH WITHIN 2 INDEX SCORES OF THRESHOLD

			2SLS MODEL		
	Baseline mean for men	Baseline mean for women	Men	Women	Difference
			(1)	(2)	(3)
Any IGA (1=Yes)	0.774	0.518	0.122	0.195*	0.073
	(0.418)	(0.500)	(0.132)	(0.118)	(0.137)
Any non-farm IGA (1=Yes)	0.471	0.195	0.100	0.399***	0.299***
	(0.499)	(0.396)	(0.138)	(0.123)	(0.143)
Trade-specific IGA (1=Yes)	0.499	0.396	0.221*	0.502***	0.281***
	(0.295)	(0.113)	(0.129)	(0.116)	(0.134)
Hours worked in past month	107.772	46.887	40.19	74.56***	34.37
	(99.126)	(70.525)	(26.44)	(23.65)	(27.42)
Total monthly earnings (NRs)	2137.947	774.683	1161	2,634***	1473
	(2539.479)	(1796.025)	(902.8)	(824.0)	(945.7)
Logged earnings	4.796	2.428	0.684	3.957***	3.273***
	(3.917)	(3.476)	(1.271)	(1.160)	(1.331)
Earnings > 3000 NRs. (1=Yes)	0.350	0.098	0.113	0.413***	0.3***
	(0.477)	(0.297)	(0.134)	(0.120)	(0.139)
Clustered Standard Errors			Yes	Yes	Yes

Notes: Standard errors (reported in brackets) clustered at the event level where possible.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 11. ESTIMATES BY BANDWIDTH CHOICE

	2SLS Estimate				
	2 Index Scores	3 Index Scores	4 Index Scores	5 Index Scores	10 Index Scores
	(1)	(2)	(3)	(4)	(5)
Full Sample (pooled across all training types)					
Any non-farm IGA	0.278**	0.265**	0.237**	0.288***	0.281***
	(0.110)	(0.104)	(0.0957)	(0.0888)	(0.0831)
Trade-specific IGA	0.387***	0.350***	0.307***	0.322***	0.319***
	(0.103)	(0.0978)	(0.0906)	(0.0849)	(0.0789)
Monthly earnings (NRs)	2,169***	2,121***	1,978***	1,752***	2,080***
	(739.9)	(697.9)	(626.7)	(575.3)	(540.5)
Earnings>3000	0.308***	0.254**	0.204**	0.187**	0.245***
	(0.107)	(0.102)	(0.0934)	(0.0866)	(0.0806)
Clustered standard errors (by course)	Yes	Yes	Yes	Yes	Yes

Notes: No poultry technician trainings were included in the 2011 sample.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

APPENDIX A

TABLE A1. EMPLOYMENT, 2010-2012 POOLED COHORTS, BANDWIDTH WITHIN 3 INDEX SCORES

Outcome	Any IGA (1=Yes)	Any non-farm IGA (1=Yes)	Trade-specific IGA (1=Yes)	Hours worked in past month	Earnings	Logged earnings	Earnings > 3000 NRs. (1=Yes)
2SLS Estimate	0.114	0.265**	0.350***	47.85**	2,121***	2.049**	0.254**
	(0.0991)	(0.104)	(0.0978)	(19.79)	(697.9)	(0.967)	(0.102)
First-stage F-statistic	158.3	158.3	158.3	158.3	151.4	151.4	158.3
Baseline mean	0.612	0.296	0.18	69.261	1271.542	3.291	0.19
	(0.487)	(0.457)	(0.384)	(87.273)	(2197.669)	(3.817)	(0.393)
Observations	1777	1777	1777	1777	1777	1777	1777

Notes: Bandwidth is within 2 index scores
Standard errors (reported in brackets) clustered at the event level where possible.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A2. EMPLOYMENT BY TRADE, 2010-2012 COHORTS, BANDWIDTH WITHIN 3 INDEX SCORES

	2SLS Estimate			
	any nonfarm IGA	trade-specific IGA	monthly earnings (NRs)	earnings > 3000
	(1)	(2)	(3)	(4)
Full Sample (pooled across all training types)				
training: Food prep/Hospitality (N=265)	-0.0384	0.186	2210	-2.200
	(0.396)	(0.391)	(2752)	(3.816)
training: Electrician & Electronics (N=641)	0.221	0.496***	2241**	0.0757
	(0.159)	(0.157)	(1081)	(0.164)
training: Handicraft & Incense stick making (N=235)	0.333	0.367	4119**	0.434
	(0.270)	(0.266)	(1822)	(0.278)
training: Construction (N=1128)	-0.117	0.165	423.4	0.129
	(0.128)	(0.126)	(890.3)	(0.132)
training: Beautician/Barber (N=239)	0.262	0.835***	4241***	0.584***
	(0.219)	(0.216)	(1474)	(0.226)
training: Weaving/Tailoring/Garment Making (N=1461)	0.276**	0.484***	2460**	0.366**
	(0.141)	(0.138)	(979.5)	(0.145)
Clustered standard errors (by event)	Yes	Yes	Yes	Yes

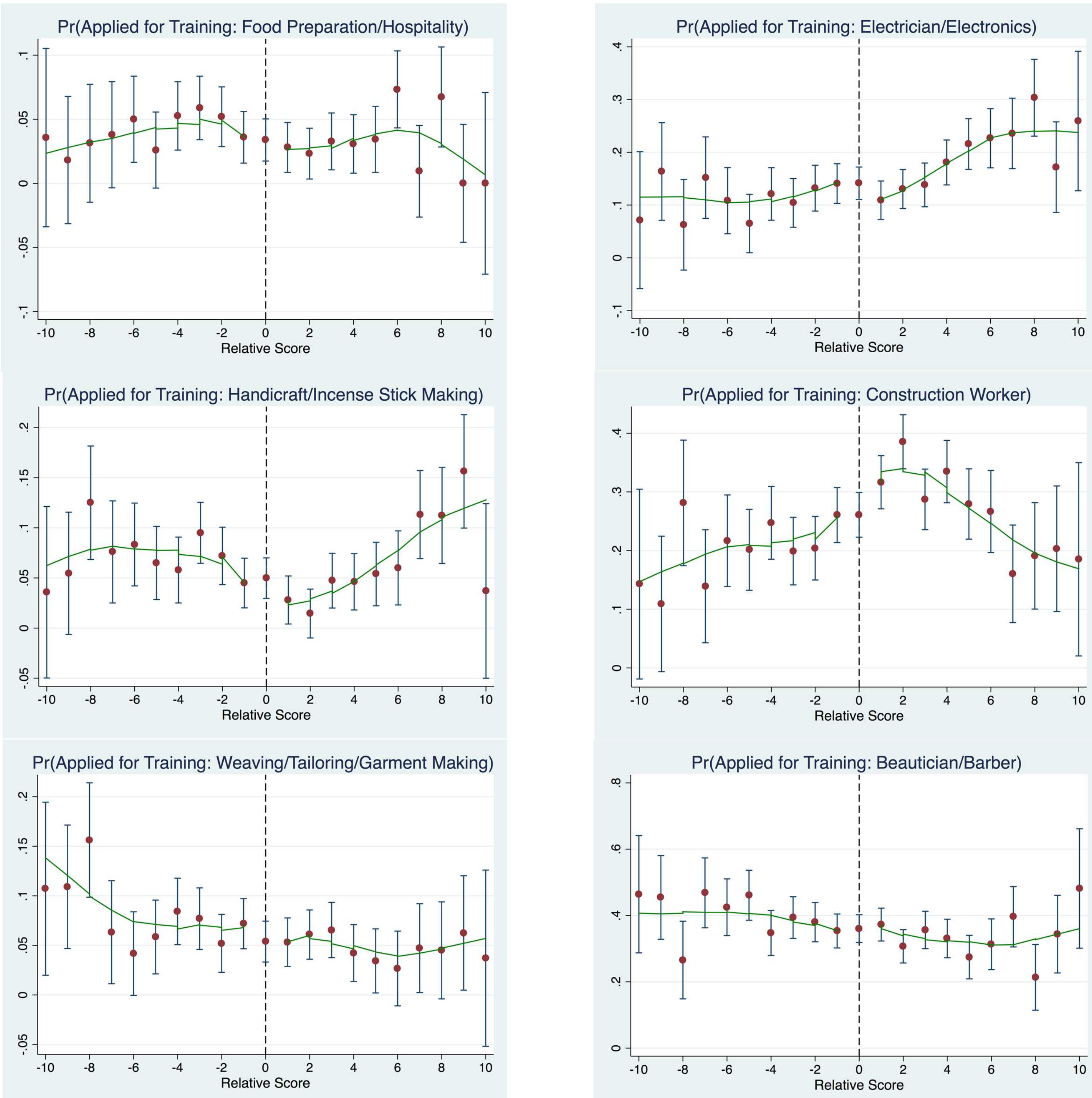
Notes: No poultry technician trainings were included in the 2011 sample.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A3. EMPLOYMENT OUTCOMES, BY GENDER, 2010-2012 POOLED COHORTS, BANDWIDTH WITHIN 3 INDEX SCORES

			2SLS MODEL	
	Baseline mean for men	Baseline mean for women	Men	Women
			(1)	(2)
'es)	0.774	0.518	0.077	0.158
	(0.418)	(0.500)	(0.117)	(0.111)
IGA (1=Yes)	0.471	0.195	0.081	0.406***
	(0.499)	(0.396)	(0.123)	(0.116)
IGA (1=Yes)	0.499	0.396	0.195*	0.467***
	(0.295)	(0.113)	(0.116)	(0.110)
in past month	107.772	46.887	24.28	62.26***
	(99.126)	(70.525)	(23.42)	(22.15)
earnings (NRs)	2137.947	774.683	903	2,653***
	(2539.479)	(1796.025)	(803.1)	(774.1)
gs	4.796	2.428	-0.109	3.598***
	(3.917)	(3.476)	(1.132)	(1.092)
00 NRs. (1=Yes)	0.350	0.098	0.054	0.361***
	(0.477)	(0.297)	(0.121)	(0.114)
dard Errors			Yes	Yes

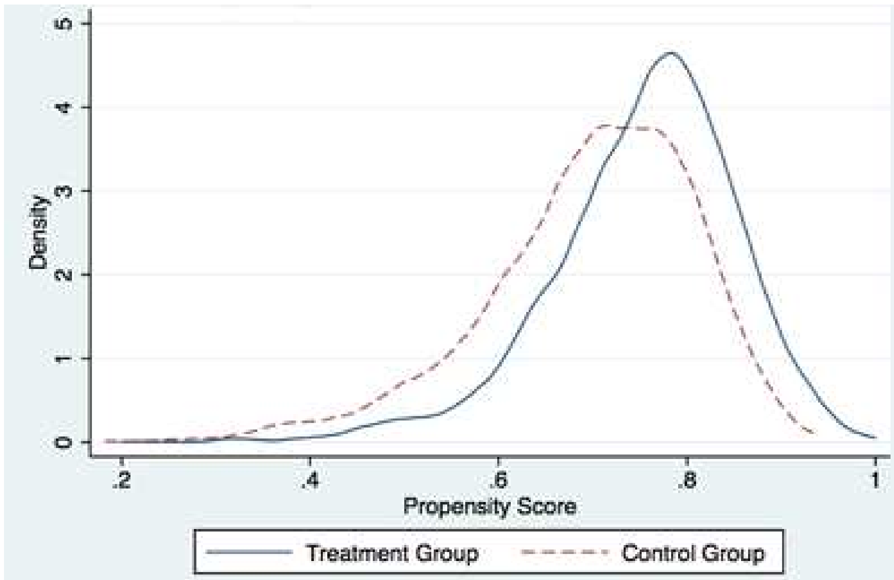
rrors (reported in brackets) clustered at the event level where possible.
the 1 percent level.
he 5 percent level.
e 10 percent level.

FIGURE A1: CONTINUITY OF INTERACTION VARIABLES AT CUT-OFF



APPENDIX B

FIGURE B1. DISTRIBUTION OF ESTIMATED PROPENSITY SCORES FOR 2010-2012 POOLED COHORTS (ITT)



Notes: Propensity Score Distributions (Baseline ITT)

TABLE B1. BASELINE BALANCING TESTS 2010-2012 POOLED (ITT), FULL SAMPLE

	Control	Treatment	Difference	p-value	N
Demographics					
Female (%)	0.640	0.630	-0.010	0.610	4101
AGEI (women aged 16-24) (%)	0.319	0.336	0.017	0.350	4101
Dalit (%)	0.090	0.077	-0.012	0.365	4037
Janajati (%)	0.421	0.468	0.048**	0.024	4037
Muslim (%)	0.017	0.025	0.008	0.269	4037
Age	24.537	24.242	-0.294	0.249	4101
Currently Married (%)	0.580	0.594	0.014	0.463	4101
Any Children (%)	0.505	0.526	0.021	0.248	4101
Completed SLC (10th grade) (%)	0.163	0.105	-0.059***	0.000	4101
Employment					
Any IGA in past month (%)	0.594	0.619	0.025	0.182	4101
Any non-farm IGA in past month (%)	0.266	0.307	0.041**	0.012	4101
Earnings in past month (NRs)	1201.970	1295.522	93.552	0.285	4069
Earnings > 3000 in past month (%)	0.172	0.197	0.025*	0.094	4101
Trade-specific IGA in past month (%)	0.154	0.189	0.035**	0.014	4101
Hours worked past month	62.774	71.502	8.728***	0.008	4101
Empowerment					
Any savings (%)	0.585	0.604	0.019	0.311	4080
Total Cash Savings (NRs)	3114.676	3177.379	62.703	0.832	4080

Notes: This table reports average values for treatment and control groups, with p-value of a Student's t-test for equality of means between the two groups. The tests are conducted on the panel sample (those interviewed at baseline and follow-up). Standard errors are clustered by training course. "ITT" indicates that treatment is defined as having a score that qualifies the respondent for an EF training course.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B2. TAKE-UP OF EF TRAINING (YEAR AFTER BASELINE SURVEY)

	Participated in an EF training course		**Did not participate in an EF training course**	
	Number	Percent	Number	Percent
2010 Cohort (N=1372)				
Assigned to Treatment (N=1040)	671	64.52%	369	35.48%
Assigned to Control (N=332)	86	25.90%	246	74.10%
2011 Cohort (N=1415)				
Assigned to Treatment (N=1110)	826	74.41%	284	25.59%
Assigned to Control (N=305)	110	36.07%	195	63.93%
2012 Cohort (N=1306)				
Assigned to Treatment (N=889)	597	67.15%	292	32.85%
Assigned to Control (N=417)	127	30.46%	290	69.54%

Notes: There are four individuals from the 2011 cohort and five individuals from 2010 whose status in the EF database is unknown. For these individuals, we rely on the respondent's self-report of whether they took an EF training in the past year for the ATT results. The table only includes those individuals who were surveyed for the first follow-up.

TABLE B3. FIRST STAGE PROPENSITY SCORE ESTIMATES

Dependent variable:	**Treat (ITT)**
Age of applicant	0.000
	(0.009)
Sex of applicant (1=Female)	0.056
	(0.094)
Education level of applicant	-0.010
	(0.010)
Education of household head	-0.003
	(0.006)
Household size	0.007
	(0.007)
Married (1=Yes)	-0.038
	(0.074)
Has child (1=Yes)	0.260**
	(0.103)
Number of children	-0.079**
	(0.039)
Any IGA (1= Yes)	0.057
	(0.072)
Zero earnings (1=Yes)	-0.075
	(0.069)
Janajati (1=Yes)	0.031
	(0.443)
Dalit (1=Yes)	0.436
	(0.690)
Muslim (1=Yes)	4.127***
	(0.357)
Analytical Ability (0-5)	-0.011
	(0.019)
Entrepreneurial score (0-32)	-0.003
	(0.004)
Financial literacy (1=Yes)	-0.047
	(0.050)
N	4449
Psuedo R^2	0.050

Notes: Standard errors (reported in brackets) clustered by training course. "Treat (ITT)" equals 1 if individual qualified for a training course and 0 otherwise. Other independent variables (not shown): district and T&E provider fixed effects, training-type categories, quintiles of household wealth. All variables measured at baseline. Although baseline data were collected on 4,677 individuals, incomplete data on ethnicity reduces the number of observations to 4,449.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B4. EMPLOYMENT (ITT), 2010-2012 POOLED COHORTS

	Baseline mean	OLS (1)	IPSW (2)	NN (3)
Any IGA (1=Yes)	0.612	0.071***	0.093***	0.070***
	[0.487]	(0.022)	(0.022)	(0.020)
Any non-farm IGA (1=Yes)	0.296	0.149***	0.160***	0.150***
	[0.457]	(0.023)	(0.024)	(0.021)
Trade-specific IGA (1=Yes)	0.18	0.182***	0.187***	0.184***
	[0.384]	(0.023)	(0.025)	(0.020)
Hours worked in past month	69.261	18.740***	21.130***	19.014***
	[87.273]	(3.890)	(4.148)	(3.940)
Earnings	1271.542	856.087***	921.323***	850.880***
	[2197.669]	(152.941)	(159.517)	(135.139)
Logged earnings	3.291	0.957***	1.209***	0.975***
	[3.817]	(0.191)	(0.203)	(0.173)
Earnings > 3000 NRs. (1=Yes)	0.19	0.130***	0.140***	0.131***
	[0.393]	(0.021)	(0.022)	(0.020)
Clustered Standard Errors		Yes	Yes	No

Notes: All columns report difference-in-difference estimates. "ITT" indicates that everyone whose score qualified them for a given training event is included in the "treatment" group.
Standard errors (reported in brackets) clustered at the event level where possible. Self-employment and location of work were not asked in 2010.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B5. EMPLOYMENT BY TRADE (ITT), 2010-2012 COHORTS

	Pooled 2010-2012 Cohorts IPSW Model (ITT Effects)			
	any nonfarm IGA	trade-specific IGA	monthly earnings (NRs)	earnings > 3000
	(1)	(2)	(3)	(4)
Full Sample (pooled across all training types)	0.160***	0.187***	921.323***	0.140***
	(0.024)	(0.025)	(159.517)	(0.022)
training: Farming (N=92)	0.155*	-0.059	1167.151	0.081
	(0.081)	(0.104)	(1000.983)	(0.169)
training: Poultry Technician (N=41)	0.226	0.342***	1139.704	0.189
	(0.173)	(0.099)	(969.082)	(0.145)
training: Food prep/Hospitality (N=265)	-0.057	0.007	-965.418	-0.146
	(0.096)	(0.064)	(1048.109)	(0.095)
training: Electrician & Electronics (N=641)	0.187***	0.258***	1282.843***	0.160***
	(0.044)	(0.058)	(359.255)	(0.054)
training: Handicraft & Incense stick making (N=235)	0.107	0.207***	967.311*	0.129
	(0.082)	(0.075)	(524.717)	(0.094)
training: Construction (N=1128)	0.067	0.100*	509.836	0.089**
	(0.054)	(0.058)	(322.866)	(0.038)
training: Beautician/Barber (N=239)	0.247***	0.402***	1529.259***	0.241***
	(0.094)	(0.089)	(533.151)	(0.078)
training: Weaving/Tailoring/Garment (N=1461)	0.249***	0.222***	1185.755***	0.196***
	(0.038)	(0.037)	(233.399)	(0.035)
Clustered standard errors (by event)	Yes	Yes	Yes	Yes

Notes: No poultry technician trainings were included in the 2011 sample.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B7. EMPLOYMENT OUTCOMES (ITT), BY GENDER, 2010-2012 POOLED COHORTS

			IPSW MODEL		
	Baseline mean for men	Baseline mean for women	Men	Women	Difference
			(1)	(2)	(3)
Any IGA (1=Yes)	0.774	0.518	0.025	0.130***	-0.106**
	[0.418]	[0.500]	(0.035)	(0.028)	(0.045)
Any non-farm IGA (1=Yes)	0.471	0.195	0.105**	0.192***	-0.087*
	[0.499]	[0.396]	(0.044)	(0.028)	(0.051)
Trade-specific IGA (1=Yes)	[0.499]	[0.396]	0.147***	0.209***	-0.062
	0.295	0.113	(0.046)	(0.028)	(0.053)
Hours worked in past month	107.772	46.887	11.564	26.287***	-14.723
	[99.126]	[70.525]	(8.796)	(4.242)	(9.795)
Total monthly earnings (NRs)	2137.947	774.683	681.698**	1036.088***	-354.390
	[2539.479]	[1796.025]	(300.488)	(173.214)	(341.802)
Logged earnings	4.796	2.428	0.281	1.688***	-1.407***
	[3.917]	[3.476]	(0.341)	(0.237)	(0.414)
Earnings > 3000 NRs. (1=Yes)	0.350	0.098	0.091**	0.166***	-0.075
	[0.477]	[0.297]	(0.039)	(0.027)	(0.047)
Clustered Standard Errors			Yes	Yes	Yes

Notes: Standard errors (reported in brackets) clustered at the event level where possible.
Younger women (aged 16 to 24) compared to older women (age 25 to 35).
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B8. EMPLOYMENT (ITT), DISAGGREGATED BY WOMEN, 2010-2012 COHORTS

			IPSW MODEL		
	Baseline mean for young women	Baseline mean for women	Young women	Older women	Difference
			(1)	(2)	(3)
Any IGA (1=Yes)	0.5	0.543	0.140***	0.127***	0.013
	[0.500]	[0.498]	(0.040)	(0.041_	(0.058)
Any non-farm IGA (1=Yes)	0.168	0.225	0.196***	0.187***	0.009
	[0.374]	[0.418]	(0.041)	(0.042)	(0.062)
Trade-specific IGA (1=Yes)	0.096	0.131	0.213***	0.204***	0.010
	[0.295]	[0.338]	(0.036)	(0.039)	(0.048)
Hours worked in past month	39.569	55.560	25.348***	27.881***	-2.533
	[62.475]	[78.058]	(5.746)	(6.83)	(9.259)
Total monthly earnings (NRs)	560.537	1026.533	834.168***	1283.426***	-449.259
	[1438.980]	[2113.341]	(183.918)	(283.993)	(320.888)
Logged earnings	2.063	2.857	1.633***	1.791***	-0.158
	[3.264]	[3.665]	(0.329)	(0.366)	(0.505)
Earnings > 3000 NRs. (1=Yes)	0.071	0.131	0.144***	0.192***	-0.048
	[0.256]	[0.337]	(0.030)	(0.043)	(0.051)
Clustered Standard Errors			Yes	Yes	Yes

Notes: Standard errors (reported in brackets) clustered at the event level where possible.
Younger women (aged 16 to 24) compared to older women (age 25 to 35).
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B9: BASELINE BALANCE TESTS 2010-2012 POOLED COHORTS (ATT)

	Control	Treatment	Difference	p-value	N
Demographics					
Female (%)	0.633	0.632	-0.000	0.994	4101
AGEI (women aged 16-24) (%)	0.333	0.331	-0.002	0.884	4101
Dalit (%)	0.097	0.069	-0.027**	0.028	4037
Janajati (%)	0.413	0.486	0.073***	0.000	4037
Muslim (%)	0.019	0.026	0.007*	0.062	4037
Age	24.389	24.268	-0.121	0.593	4101
Currently Married (%)	0.585	0.595	0.010	0.577	4101
Any Children (%)	0.509	0.529	0.020	0.272	4101
Completed SLC (10th grade) (%)	0.147	0.101	-0.046***	0.000	4101
Employment					
Any IGA in past month (%)	0.593	0.625	0.032*	0.082	4101
Any non-farm IGA in past month (%)	0.283	0.306	0.023	0.142	4101
Earnings in past month (NRs)	1258.539	1280.494	21.955	0.788	4069
Earnings > 3000 in past month (%)	0.185	0.194	0.010	0.490	4101
Trade-specific IGA in past month (%)	0.173	0.185	0.011	0.397	4101
Hours worked past month	66.292	71.308	5.016	0.128	4101
Empowerment					
Any savings (%)	0.580	0.612	0.032*	0.079	4080
Total Cash Savings (NRs)	3246.505	3102.511	-143.994	0.608	4080

Notes: This table reports average values for treatment and control groups, with p-value of a Student's t-test for equality of means between the two groups. The tests are conducted on the panel sample (those interviewed at baseline and follow-up). Standard errors are clustered by training course. "ITT" indicates that treatment is defined as having a score that qualifies the respondent for an EF training course.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B10: FIRST STAGE PROPENSITY SCORES (ATT)

Dependent variable:	**TREAT**
Age of applicant	0.001
	(0.009)
Sex of applicant (1=Female)	0.026
	(0.086)
Education level of applicant	0.005
	(0.009)
Education of hh head	-0.001
	(0.006)
Household size	0.008
	(0.006)
Married (1=Yes)	-0.055
	(0.062)
Has child (1=Yes)	0.291***
	(0.086)
Number of children	-0.113***
	(0.034)
Any IGA (1= Yes)	0.117*
	(0.069)
Zero earnings (1=Yes)	-0.034
	(0.062)
Janajati (1=Yes)	0.483
	(0.661)
Dalit (1=Yes)	0.146
	(0.798)
Muslim (1=Yes)	4.865***
	(0.370)
Analytical Ability (0-5)	0.028
	(0.018)
Entrepreneurial score (0-32)	-0.004
	(0.004)
Financial literacy (1=Yes)	-0.053
	(0.050)
N	4490
Psuedo R^2	0.071

Notes: Standard errors (reported in brackets) clustered by training course. "Treat (ATT)" equals 1 if individual participated in a training course and 0 otherwise. Other independent variables (not shown): district and T&E provider fixed effects, training-type categories, quintiles of household wealth. All variables measured at baseline. Although baseline data were collected on 4,677 individuals, incomplete data on ethnicity reduces the number of observations to 4,449.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B11: EMPLOYMENT (ATT), 2010-2012 COHORTS

	Baseline mean	OLS (1)	IPSW (2)	NN (3)
Any IGA (1=Yes)	0.612	0.085***	0.123***	0.089***
	[0.487]	(0.021)	(0.021)	(0.018)
Any non-farm IGA (1=Yes)	0.296	0.203***	0.203***	0.206***
	[0.457]	(0.024)	(0.025)	(0.019)
Trade-specific IGA (1=Yes)	0.18	0.244***	0.233***	0.243***
	[0.384]	(0.024)	(0.024)	(0.017)
Hours worked in past month	69.261	29.581***	31.545***	30.484***
	[87.273]	(4.092)	(4.171)	(3.475)
Earnings	1271.542	976.240***	1099.759***	1018.001***
	[2197.669]	(135.081)	(131.653)	(119.605)
Logged earnings	3.291	1.392***	1.554***	1.432***
	[3.817]	(0.195)	(0.195)	(0.152)
Earnings > 3000 NRs. (1=Yes)	0.19	0.159***	0.168***	0.161***
	[0.393]	(0.022)	(0.021)	(0.018)
Clustered Standard Errors		Yes	Yes	No

Notes: All columns report difference-in-difference estimates. "ITT" indicates that everyone whose score qualified them for a given training event is included in the "treatment" group.
Standard errors (reported in brackets) clustered at the event level where possible.
Self-employment and location of work were not asked in 2010.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B12: EMPLOYMENT 2010-2012 POOLED (ATT IPSW MODEL)

	any nonfarm IGA	trade-specific IGA	monthly earnings (NRs)	earnings > 3000
	(1)	(2)	(3)	(4)
Full Sample (pooled across all training types)	0.203***	0.233***	1099.759***	0.168***
	(0.025)	(0.024)	(131.653)	(0.021)
training: Farming (N=92)	0.139**	-0.044	529.452	0.081
	(0.068)	(0.083)	(547.819)	(0.094)
training:Poultry (N=41)	0.036***	0.184*	-707.959	-0.085**
	(0.011)	(0.110)	(442.371)	(0.040)
training:Food prep/Hospitality (N=265)	0.164	0.139*	1596.182***	0.103
	(0.139)	(0.080)	(581.698)	(0.079)
training:Electrician & Electronics (N=641)	0.191***	0.353***	884.423**	0.140***
	(0.064)	(0.054)	(373.850)	(0.047)
training:Handicraft & Incense stick making (N=235)	0.047	0.162**	1594.242***	0.178**
	(0.081)	(0.073)	(450.527)	(0.073)
training: Construction (N=1128)	0.082**	0.079*	702.008**	0.099**
	(0.040)	(0.041)	(298.111)	(0.045)
training: Beautician/Barber (N=239)	0.459***	0.501***	1966.409***	0.282***
	(0.108)	(0.114)	(402.549)	(0.069)
training: Weaving/Tailoring/Garment Making (N=1461)	0.303***	0.306***	1272.605***	0.235***
	(0.039)	(0.038)	(180.795)	(0.034)
Clustered standard errors (by event)	Yes	Yes	Yes	Yes

Note: No poultry technician trainings were included in the 2011 or 2012 samples.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B13: EMPLOYMENT (ATT-IPSW MODEL), BY GENDER, 2010-2012 COHORTS

	Baseline mean for men	Baseline mean for women	Men	Women	Difference between men and women
	[Std Dev]	[Std Dev]	(1)	(2)	(3)
Any IGA (1=Yes)	0.774	0.518	0.091***	0.143***	-0.051
	[0.418]	[0.500]	(0.030)	(0.028)	(0.040)
Any non-farm IGA (1=Yes)	0.471	0.195	0.128***	0.247***	-0.120**
	[0.499]	[0.396]	(0.041)	(0.030)	(0.050)
Trade-specific IGA (1=Yes)	0.295	0.113	0.170***	0.269***	-0.098**
	[0.456]	[0.317]	(0.038)	(0.030)	(0.048)
Hours worked in past month	107.772	46.887	23.887***	35.808***	-11.921
	[99.126]	[70.525]	(7.103)	(4.758)	(8.147)
Total monthly earnings (NRs)	2137.947	774.683	898.782***	1185.170***	-286.387
	[2539.479]	[1796.025]	(255.902)	(141.973)	(286.878)
Logged earnings	4.796	2.428	0.582*	2.072***	-1.490***
	[3.917]	[3.476]	(0.300)	(0.236)	(0.377)
Earnings > 3000 NRs. (1=Yes)	0.350	0.098	0.115***	0.197***	-0.081*
	[0.477]	[0.297]	(0.035)	(0.025)	(0.042)
Clustered Standard Errors			Yes	Yes	Yes

Notes: Standard errors (reported in brackets) clustered at the event level where possible.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B14: EMPLOYMENT (ATT-IPSW MODEL), FOR WOMEN, 2010-2012 COHORTS

	Baseline mean for young women	Baseline mean for women	Young women	Older women	Difference
	[Std Dev]	[Std Dev]	(1)	(2)	(3)
Any IGA (1=Yes)	0.500	0.543	0.137***	0.159***	-0.022
	[0.500]	[0.498]	(0.039)	(0.037)	(0.051)
Any non-farm IGA (1=Yes)	0.168	0.225	0.246***	0.252***	-0.006
	[0.374]	[0.418]	(0.039)	(0.041)	(0.052)
Trade-specific IGA (1=Yes)	0.096	0.131	0.265***	0.276***	-0.011
	[0.295]	[0.338]	(0.036)	(0.037)	(0.042)
Hours worked in past month	39.569	55.560	30.988***	42.010***	-11.022
	[62.475]	[78.058]	(6.019)	(7.538)	(9.525)
Total monthly earnings (NRs)	560.537	1026.533	1063.272***	1343.550***	-280.278
	[1438.980]	[2113.341]	(172.773)	(218.104)	(263.357)
Logged earnings	2.063	2.857	2.028***	2.187***	-0.159
	[3.264]	[3.665]	(0.301)	(0.325)	(0.406)
Earnings > 3000 NRs. (1=Yes)	0.071	0.131	0.165***	0.235***	-0.071*
	[0.256]	[0.337]	(0.030)	(0.036)	(0.040)
Clustered Standard Errors			Yes	Yes	Yes

Notes: Standard errors (reported in brackets) clustered at the event level where possible.
Younger women (aged 16 to 24) compared to older women (age 25 to 35).
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.