The Role of Trainee Selection in the Effectiveness of Vocational Training: Evidence from a Randomized Controlled Trial in Nepal*

Shyamal Chowdhury[1], Syed Hasan[2], and Uttam Sharma[3]

[1]School of Economics, The University of Sydney, Australia and IZA, Germany
[2]School of Economics and Finance, Massey University, New Zealand
[3]Institute for Social and Environmental Research, Nepal (ISER-N) and University of Michigan, USA

September 15, 2024

Abstract

Based on a randomized controlled trial conducted on extremely poor youths in Nepal, we report the impact of a vocational training program that offered three-months training combined with incentives for trainers tied to trainees' success. Furthermore, to mimic practices in the field, a component of the program allowed trainers to select trainees from eligible applicants. For the trainees that were randomly selected, nine months after program completion, we found no significant effect of the training on the outcomes except for employment prospects. However, we observed some improved outcomes for the trainees selected by trainers. These findings are consistent with the observed pattern of better outcomes when program implementers non-randomly assign treatments. We also found no significant effect of selection. Thus our investigation suggets that trainee selection in vocational training programs can provide a better outcome in low-income countries.

JEL-Classification: L25, L26, L53, M53, O12

Keywords: Vocational training, job training, employment training, impact evaluation, RCT

*The IRB approval was obtained from the University of Sydney's Human Research Ethics Committee (Project No.: 2013/980); ISER-N IRB Approval (for interview of the trainers) No: A-014/2080/081, Date: August 25, 2023. The trial was post-registered at the American Economic Association's registry for randomized controlled trials, ID: 7207 (https://www.socialscienceregistry.org/trials/7207). Financial support by the DFID (now FCDO) and the Employment Fund is greatly acknowledged. We thank David McKenzie, Christopher Woodruff and participants in the 2023 Australasian Development Economics Workshop (ADEW), Massey University Economics seminar and University of Minnesota Applied Economics seminar for commenting on an earlier draft. All correspondence to Shyamal Chowdhury, School of Economics, University of Sydney, Australia. email: shyamal.chowdhury@sydney.edu.au.

1. Introduction

Enhancing the capacity of poor youths through vocational training is often prescribed as a solution to poverty and unemployment in developing countries. As a result, often with the support of international donors, governments in those economies invest heavily in vocational training (Acevedo et al., 2020; Doerr, 2022; Katz et al., 2022; Barrera-Osorio et al., 2023). This article re-examines the effectiveness of such training by providing evidence from a carefully designed incentive-based program that offered vocational training in Nepal in a randomized controlled trial (RCT) setting. Evaluating the program we reconfirm that vocational training has a limited short run effect on economic outcomes. We subsequently explore one potential way to improve the vocational training program design and conclude that trainer selection of trainees can be effective in this regard.

We study the short-run impact of Nepal's youth vocational training program "Path to Prosperity" for four specific reasons. First, Nepal relies heavily on vocational training programs as a strategy to reduce its poverty, so causal evidence of their effectiveness can be helpful in poverty reduction (Employment Fund, 2013; Asian Development Bank, 2017). However, no randomized assignment-based research design has been employed to evaluate their effectiveness. Second, the training incorporate an incentive mechanism in which the training providers receive part of the payment only if the trainees are employed within the first six months of training completion. The mechanism is likely to make the trainees receive the best possible training.

Third, the length of the training is three months, reasonable compared to many other similar training programs (e.g., Barrera-Osorio et al., 2023), and thus more likely to detect the impact of the training, if any. Shorter training, even when highly effective, can generate low benefits that are difficult to detect statistically.[1] A fourth advantage is the opportunity to examine the case when trainers select the trainees, as we have convinced the policymakers to retain randomly selected (about) half of the eligible candidates for a trial for that purpose. We compare the outcome of the trainer selected group with that of the randomly selected trainees to find whether a better outcome–conditional (on the trainer selection) average treatment effect (CATE)–can be achieved.

[1] Relying on a meta-analysis with 200 recent studies, Card et al. (2018) find training to have modest positive effects only in the long-run. Since training costs are usually low, the implied returns are higher than those in education.

The results in this article highlight that vocational training may generate more short-run benefits when trainers select the trainees, as trainers may have important information on the unobservable trainee characteristics that allow them in finding suitable candidates with higher return. Previous studies on the effectiveness of training programs in developing countries find both positive (e.g., Maitra & Mani, 2017; Alfonsi et al., 2020; Das, 2021) and small or null (e.g., Card et al., 2011; Cho et al., 2013; Blattman et al., 2020) effects.[2] Our study can confirm the previous findings and, by focusing on program design, suggest an effective way to provide higher benefits.[3]

Random program assignment can provide a more reliable causal effect of training. Yet, non-random assignments are common in training programs. So, it is worthwhile to examine the consequence of such selection on the outcome.[4] In fact, while most of the vocational training for unemployed youth were ineffective (see, Heckman et al., 1999; McKenzie, 2017; Agarwal & Mani, 2024), some studies, mostly non-experimental, find the contrary (e.g. Chakravarty et al., 2019; Van den Berg & Vikström, 2022). For instance, using a regression-discontinuity design, Chakravarty et al. (2019) find vocational training to raise non-farm employment and earnings of Nepalese youth.[5]

We address this issue by using a setting that allow us to examine whether trainer-selected trainees do better than randomly-selected trainees from an identical trainee pool.[6,7] This is particularly interesting as the program uses an incentive-based payment system to motivate trainers to select candidates with a higher success potential. Moreover, the program has reasonable training

[2]Vocational training in developing countries vary with regard to the target population, training type, length and contents and the provision of certification (Agarwal & Mani, 2024).

[3]The impacts can differ between the short and long run. For example, randomly provided small unsupervised grants to young adults in Uganda's conflict-affected north increase their business assets, work hours, and earnings, but those benefits disappear after nine years (Blattman et al., 2013, 2020). In contrast, large effects of training on formal employment and earning persist in the long run (Attanasio et al., 2011, 2017). Agarwal & Mani (2024), on the other hand, find no over time differences in the impacts of skills training programs.

[4]Training providers, both public and private, often resist random assignments as they arguably can identify applicants who are most likely to benefit from such training. If program participants are selected non-randomly, the treatment and control group participants differ in observable and/or unobservable characteristics before the program's implementation. As a result, differences in outcomes between selected (by the program implementers) and not-selected participants can be wrongly attributed to the program.

[5]Variation in the estimated impact of microcredit can be considered as a classic example of the selection issue. RCT-based studies find only a modest impact of microcredit on borrowers' income growth and poverty reduction (Banerjee et al., 2015). In contrast, non-experimental studies, which are likely to suffer from selection issues, mostly found positive impacts (e.g. Pitt & Khandker, 1998; Khandker, 2005). However, researchers have not excluded the possibility that the effect of microcredit can vary among subgroups (e.g. Banerjee et al., 2018).

[6]Heckman et al. (1999) found the gains from vocational training to generally low as they target unskilled and less able individuals. Card et al. (2018) found that selection is important in matching training type with enterprise type. Rodríguez et al. (2022) found the average returns to training to vary across the unobserved ability distribution.

[7]Another option for selecting effective candidate is through providing incentive to the applicants for program participation, as young people possess valuable skills that are unobservable to employers (Abebe et al., 2021a) and application incentive improves the quality of the applicant pool (Abebe et al., 2021b).

duration, which is likely to generate detectable benefits. Evaluating the training programs in this way allows us to overcome the low statistical power issue faced by many earlier studies (McKenzie, 2017).[8]

Specifically, our study follows a two-stage procedure to examine whether outcomes improve with the trainer selection of trainees. In the first stage, the research team randomly divides the eligible applicants into two parts. In the second stage, in one part, trainees have been selected randomly to match the number of spots available. In the other part, trainers have selected whom they wanted to train, excluding the rest from getting any training. It means that, in the latter case, the selection of trainees has not been random but rather chosen by the trainers from the eligible candidates.

We examine whether vocational training benefited low-income youths in terms of employment, working hour, income, business ownership, and international migration. Our study reveal that randomly assigned training participants become 18 percentage points (pp) more likely to be employed, but other outcomes do not change significantly. In contrast, when trainers select the trainees, their employment prospects, working hours, and likelihood of international migration increase by 27 pp, 38 hours, and 7 pp, respectively.[9] The pattern is generally consistent with the use of regression adjustment, inverse probability weighting, covariate selection by LASSO, randomization inference test, and multiple hypotheses corrected p-values, and in the presence of treatment heterogeneity.

The estimated employment effect for the trainer-selected group is higher than the corresponding estimates in some previous studies on Nepal (e.g., Chakravarty et al., 2019). The effect on income also seems large in a country with high poverty incidence, as individual benefits are close to the poverty threshold.[10] Thus, this study may assist Nepal and other low-income countries by suggesting how to design vocational training and whom to target for maximising program benefits.

We further examine whether a higher impact on the trainer-selected participants can be attributed to the positive selection of the trainees. To do so, we investigate the post-training outcomes of the group who were left out by the trainers. Furthermore, we use the trainers' selection criteria to choose from the randomized control participants who are comparable to the trainer-selected participants and compared their outcomes. We have also compared the part of the randomized

[8] Simultaneously to the vocational training, we have conducted an RCT on entrepreneurship training. As the two studies belong to different strands of literature, we have not discussed the outcome of entrepreneurship training here.

[9] The coefficients are significant at the 5 percent level against a one-sided alternative.

[10] Poverty threshold in Nepal is defined by per capita consumption of NRs.3,500 (NRs. stands for Nepalese Rupees) per month in 2015 (Asian Development Bank, 2017).

control participants who would have been selected by the interviewers against those who would not. In all cases, we find no evidence of a direct effect of selection, indicating that the return is higher for the selected trainees than those who are randomly selected.

By confirming that training can be more effective when trainers select the trainees, we make important contributions to the literature on vocational training programs in developing countries. Since trainers in the program receive full payment only when trainees secure employment, they are likely to put in more effort to ensure that their graduates are employed. Additionally, trainers are likely to select trainees with a higher likelihood of success (e.g., those with greater motivation and/or capability), as they may better understand applicants' characteristics that are not observed by researchers or policymakers. The decision-making power bestowed upon them may also positively motive trainers. By comparing the magnitudes of the impacts with and without trainer selection of trainees, we can identify the contribution of trainer selection. Subsequent interviews with trainers, confirming the hypothesized selection mechanism, we join the literature on design and effectiveness of training programs. Our focus on targeting also contributes to the impact evaluation literature by comparing outcomes under alternative targeting policies.[11]

The rest of the paper proceeds as follows. Section 2 describes the program background and research design, including sampling procedures, training details, and the timeline of activities. Section 3 discusses the empirical model, data and the attrition issue. Section 4 presents the main results, including the robustness and heterogeneity checks. Policy implications of our research and costing issues are discussed in Section 5. Section 6 concludes with a discussion on justifying the potential of scaling up the program.

2. Research context and research design

2.1. Background

The flagship training program evaluated in this study is "The Skills Training and Employment Services for the Very Poor and Youth with Special Needs (Path to Prosperity)," providing vocational training to the extremely poor youths in Nepal. The program was implemented by a large Nepal-

[11] Recent studies evaluated alternative targeting policies using observable characteristics/features in the data and machine learning technique (e.g., Blumenstock et al., 2015; Aiken et al., 2022, 2023; Athey et al., 2023). We, however, evaluate the use of trainers' insights on unobservable trainees characteristics in selecting the training participants.

based NGO, the Employment Fund (EF), with financial support from the UK's Department for International Development (DFID), the Swiss Agency for Development and Cooperation (SDC), and the World Bank. The training program was a part of a larger anti-poverty initiatives aimed at stimulating microenterprise and employment opportunities for low-income people by providing vocational and entrepreneurship training to about 55,000 trainees per year.[12] The training was conducted in 23 of Nepal's 75 districts in early 2014 (Employment Fund, 2013).

2.2. The training

The training program evaluated in this study had two important features. Firstly, trainers were offered explicit incentive-based payment. Specifically, the trainers received the final 60 percent of their remuneration if the trainees became employed within three to six months after the training (Employment Fund, 2013). Secondly, unlike many vocational training programs in low-income countries, the training was more extensive, with each trainee receiving three months of training, making the effect more likely to be detected econometrically.

In the program, each trainee was trained for at least 390 hours (equivalent to three months of intensive training), of which one-third was dedicated to on-the-job training/apprenticeship-based learning. The training was exclusively offered to the extremely poor youth and focused on common occupations in Nepal, such as furniture making, handicraft manufacturing, tailoring/garment making, food catering, hospitality service, and brick-making. Excluding administrative expenditures, the training cost was around NRs.40,000 (≈US$400) per participant, with slight variations across training types and providers. Participants took the training free of charge.[13]

2.3. Research design

Our research relies on an RCT design, where the allocation into treatment and control groups from the eligible training applicants involves two steps. In the first step, applicants are randomly divided into two groups. In the second step, from the first group, a predetermined number of training participants are randomly selected based on the number of available training spots. We

[12] In 2013, EF was responsible for around 30 percent of the trainees participating in vocational and entrepreneurial training programs in Nepal (Employment Fund, 2013).

[13] We learned about the costs through personal communications. For other details, see Employment Fund (2013).

refer to this as the random treatment (RT) group. The remainder forms the random control (RC) group. The RT and RC groups together provided our *first analysis sample*.

In the second group, during the second step, trainers selected a predetermined number of trainees based on the number of the available training spots. These selected participants–referred to as the trainer treatment (TT) group–are compared against the RC group. Thus, the TT and RC groups constitute our *second analysis sample*. We also compare the outcomes of the remaining participants–the trainer control (TC) group–against the RC group. The TC and RC groups together form our *third analysis sample*. Our research is designed to estimate the effects of vocational training with trainers' incentive, further exploring whether outcomes differ when trainers select trainees—a practice common in the field.

2.4. Sampling and randomization

Applicants aged between 18 and 40 years and not enrolled in formal education at the time of the application were eligible for program participation. EF relied on the Training & Employment (T&E) providers to select applicants based on their own guidelines (Employment Fund, 2013). The total number of participants in the program was 1,036, which appears reasonable compared to previous studies detecting the effect of the training.[14] This study include the 34 vocational training events organized by EF across Nepal. Each event typically trained around 22 trainees, meaning there were about 748 training spots.

To examine the effect of selection, all program participants were ranked by a suitability score that relied on an interview by the T&E providers. The interviewers attempted to assess motivation, network, and ability of the participants related to the job that they were interested in. Participants with family members working in the same field of training were given higher scores. Applicants from disadvantaged backgrounds were also given some additional scores in the relevant categories to prioritize their inclusion. No specific scoring criteria were given to the T&E providers to ensure that the selection reflected the selectors' understanding of local situations.

Out of the total applicants selected for this research, the *first analysis sample* contains 512 applicants from the first group. Of these, 373 persons were randomly selected for participating in one of the 17 training events (RT group). The remaining 139 individuals were kept in the RC

[14] A review by Agarwal & Mani (2024) provided a list of studies on low income countries and their sample sizes.

group. The motivation for selecting a specific number of trainees was to fulfill the available training capacity. Of the retained 524 persons in the second group, solely on the basis of the suitability score, 374 applicants (TT group) were selected by the T&E providers to participate in one of the 17 training events. The TT and RC group members (from the first group) provided our *second analysis sample* of size 513. The remaining 150 applicants (TC) from the second group and the RC group from the first group constituted in our *third analysis sample* of size 289 (Table 1).[15]

[Table 1]

2.5. Study Timeline and Data Collection

The baseline survey was conducted from March to early April of 2014, before the program was implemented. The training programs started in late April and concluded in early July of 2014. The endline survey began in March 2015, nine months after the training ended. Data collection was halted temporarily due to a major earthquake in Nepal on 25 April 2015. The survey resumed on 28 May 2015 and was completed on 22 July 2015. For those living outside Nepal, whenever possible, phone interviews were conducted. Both rounds of the survey employed a similar set of questionnaire, although a shorter questionnaire was used for those interviewed through phone.[16]

2.6. Data

We selected a specialized survey company, Nielsen, through a competitive bidding process to collect data from the study participants. Nielsen, a survey firm with a proven track record, collected baseline and endline information for all study participants who were available at the time of the interview.

Nielsen collected information on the following outcome variables: i) whether the applicant was gainfully employed, ii) total hours worked in the last month, iii) income in the last month, iv) monthly income working for oneself in the last month, v) whether the person owns a business,

[15] Event-wise scores were higher for three TC group members compared to their TT group counterparts. This happened because, instead of the one with better score, the next best participant from the wait list entered into the TT group. Dropping them does not affect our conclusions, so we ignored the issue in our analysis. Furthermore, two RT and four TT group members had missing suitability scores. We dropped them from our analysis.

[16] We conducted the follow-up survey at least three months after the verification of employment, which was the basis for incentive payment. We did so as, by our research design, a significant portion of training providers' earnings relied on trainees' post-training employment. As a result, the training providers had incentives to collude with firms to hire their trainees for a brief period.

and vi) whether the applicant has migrated overseas.[17] Those outcome variables are commonly employed in the studies on vocational training (e.g., Cho & Honorati, 2014; Blattman & Ralston, 2015). They were considered important to indicate the intervention's effectiveness (McKenzie, 2017; Agarwal & Mani, 2024), and so we chose them as the primary outcome variables in our analysis.

Nielsen also collected information for another set of outcome variables similar to the primary outcome variables: i) gainfully employed (including home cultivation), ii) average daily hours worked, iii) internal migration, iv) formal family business, and v) other family members' income. Since they have limited usefulness in explaining/complementing the main sets of results, we occasionally discussed them but included the results in the appendix.

The information related to the control variables were age, sex, years of education, marital status, and caste. The selection of the control variables, important for modeling training outcome and thus reducing the error variance, were based on previous studies (e.g., Hirshleifer et al., 2016; Acevedo et al., 2020; Baird et al., 2022). The continuous independent variables were converted into group dummies to make the estimates consistent, as suggested in J-PAL (2022).

2.7. Balance check and summary statistics

We examine the summary statistics for the control and outcome variables collected in the baseline survey to check whether the groups were balanced before the intervention, and thus, the setup remains valid for the unbiased estimation of the TE. Table 2 provides the means and standard errors (SEs) for all control and outcome variables organized under four groups: i) RC (column 1), ii) RT (column 2), iii) TT (column 4), and iv) TC (column 6). We also present the differences in means separately between RT, TT and TC groups and the RC group—the universal referance group in our investigations (columns 3, 5 and 7, respectively).

[Table 2]

Table 2 shows no systematic differences in the control variables between the baseline values of the RC group and each of the RT, TT and TC groups (Panel a). The *F-test* results confirm that the differences in the control variables between each pairs are not jointly significant. Looking at the

[17] We used the level values of all monetary dependent variables since using their logarithmic transformations may artificially show very high treatment effects for certain outcomes with baseline values close to zero.

outcome variables, we only find a significant difference for gainful employment and income between TT and RC, and for income between TC and RC (Panel b). The *F-test* results again show that the differences in the outcome variables between each pairs are not jointly significant.

2.8. Attrition

Of the 1,036 study participants in the baseline, 241 (23.2 percent) could not be contacted in person in the follow-up survey. Of those, 80 participants were outside Nepal, and their interviews were taken over the phone using a shorter questionnaire. The remaining 161 participants attrited during the endline, and the attrition rates were slightly higher for the RC group than the RT or TT groups but lower than the TC group (Appendix Table A.1).[18] The overall attrition rate in our study (15.5 percent) was similar to the corresponding average of the recent RCT-based studies on vocational training (15 percent). Furthermore, the take-up rate in our study was about 84 percent—higher than the median take-up rate of 70 percent for similar studies; see Agarwal & Mani (2024).

To investigate the impact of treatment assignment on attrition, we regressed attrition on group assignment (RT, TT and TC) and the control variables using our main specification (equation (1) in Section 3), as suggested in J-PAL (2022). We found that group assignments were negatively associated with attrition of study participants for the RT, TT and TC groups, but none were statistically significant at the 5 percent level (Appendix Table A.2). Furthermore, the interaction of the group assignment variables with applicant characteristics did not show any particular pattern for attrition.

Next, separately for all three analysis samples, we examined the differences in control and outcome variables between the attrited members of the RC (reference) group and the RT, TT and TC groups (separately). Results indicated no systematic differences in almost all the characteristics between the RC and the RT group (Appendix Table A.3, Panel a). We found similar results for the TT and TC group members. Also, there were no significant differences in the outcome variables of the reference (RC) group and the RT, TT and TC groups in the attired sample, except for hours worked of the TC group (Appendix Table A.3, Panel b). Importantly, the *F-test* results indicate

[18] In our endline data, the missing values for the outcome variables were distributed as follows: i) 111 for gainful employment (in which, we considered overseas applicants as gainfully employed even if we could not interview them), ii) 211 for the last month's total working hour, iii) 161 for last month's income, iv) 161 for monthly income working for oneself, v) 161 for business ownership, and vi) 111 for international migration.

insignificant differences between the RC and the RT groups, jointly either for the control variables or the outcome variables; we find similar results for the TT and TC groups. Nevertheless, to address any concern about the missing values and attrition, we conducted attrition-adjusted tests throughout the analysis to ensure that our estimates of the treatment effects (TEs) remains valid for policy.

3. Empirical method

With a randomized setting, we use the following linear regression model for our investigations:

$$y_i = \beta_0 + \beta_1 \; RT_i + \boldsymbol{\theta X_i} + \varphi_d + \varepsilon_i, \tag{1}$$

where, for each individual i, y is one of the six outcome variables discussed in Subsection 2.6. RT is a binary variable taking the value of 1 if an individual belongs to the RT group and 0 otherwise. The vector $\boldsymbol{X}$ lists baseline-level controls related to individual and household characteristics, including the baseline outcome. φ_d represents district fixed effects (FEs) while ε is a mean-zero error term. We employed a similar model to investigate the case of the TT group, in a separate analysis.

The coefficient β_1 in equation (1) captures the intention to treat (ITT) effects in our setting. It is the most policy-relevant parameter as it captures the low-compliance issue we observe in practice. With a high compliance rate, as the case is for our study, the estimate will be close to the average treatment effects (ATEs).

Next, we combine the two groups to allow for a larger sample and estimate the following regression model:

$$y_i = \beta_0 + \beta_1 \; RT_i + \beta_2 \; TT_i + \boldsymbol{\theta X_i} + \varphi_d + \varepsilon_i, \tag{2}$$

where, RT and TT are the relevant group indicators (RC is the reference group).

Additionally, to examine whether the post-training differences between the RT and TT groups are statistically significant, we estimate the following model:

$$y_i = \beta_0 + \beta_1 \; Treatment_i + \beta_2 \; Treatment_i \times TT_i + \boldsymbol{\theta X_i} + \varphi_d + \varepsilon_i, \tag{3}$$

where, the variable *Treatment* takes the value of 1 for RT or TT groups and 0 for the RC (reference) group. The coefficient β_1 indicates the effect of training on the RT group. The coefficient β_2 indicates whether the effect on the TT group is significantly higher than that of the RT group.

We follow certain norms to improve the quality and consistency of the analysis in our study. First, as suggested in Athey & Imbens (2017); Wooldridge (2021); Abadie et al. (2023), we use robust standard error to account for heteroskedasticity and clustered them at the district level to address the issue that treatment assignment is based on the available training spots in districts. Second, we set a random seed and employ 1,000 replications for bootstrapping to ensure the replicability of the results. Third, we follow the discipline's convention of using the 5 percent significance level for our hypotheses testing.

4. Results

We begin our investigation by comparing the post-intervention outcomes of the RT, TT and TC groups with the RC (reference) group. Analysis using the *first analysis sample*, indicate a statistically (and practically) significant effect on employment of the RT group (Appendix Table A.4). However, jointly for all the outcomes, the effect is not statistically significant. Analysis with the *second analysis sample*, on the other hand, shows gainful employment and business ownership to be significantly higher for the TT group. Furthermore, when we conduct an *F-test* of joint significance of all the outcome variables, we reject the null hypothesis of no difference between the groups. Comparison of TC with RC group, using the *third analysis sample*, find income of the former group to be significantly lower. However, we fail to reject the null hypothesis of no differences between TC with RC groups, jointly for all the outcome variables.

To motivate whether training outcomes can be improved by trainee selection, we proceed with separately estimating the effects on RT and TT groups compared to the RC group, using model (1). Next, to improve statistical significance in our estimates, we use model (2) and the *first analysis sample* plus the TT group observations.[19] Then, using model (3), we investigate whether the outcome has been improved significantly for the TT group compared to the RT group. Finally,

[19] The *second analysis sample* includes both the TT the RC group members. From that, we only take the former group as the latter group is already included in the *first analysis sample*.

using model (1), we investigate the changes in the TC groups to conclude whether a better outcome for the TT group can be attributed to the positive selection of the trainer-selected trainees.

(SC: I think the following paragraph can be moved to the previous section.)

Throughout our analysis, we have conducted several robustness checks of our estimates of the TEs in each part of our analysis. First, we use regression adjustment (RA) that contrasts the averages of treatment-specific predicted outcomes to estimate the TEs. This method is useful when there is a selection bias in the RCTs, which generally produces misleading results (Allcott, 2015; Słoczyński, 2022; Krauss, 2018, 2021). RA can produce the TE estimates that are robust of any potential selection bias. Second, we employ inverse probability-weighted regression adjustment (IPWRA), which uses weighted regression coefficients to compute averages of the treatment-level predicted outcomes, where the weights are the estimated inverse probabilities of being assigned to the treatment. The contrasts of these averages are used to estimate the TEs.[20] Third, we use augmented inverse-probability weighting (AIPW) with the selection of covariates using a machine learning approach, Lasso.[21] Fourth, we estimate the Lee bounds and tighten them by adding covariates, as suggested in Lee (2009) and J-PAL (2022).[22] Fifth, we examine the significance of the estimated TEs with the randomization inference method.[23]

Once we conduct our main analysis, we examine whether our conclusions remain unaltered when we use statistical significance that corrects for the case of multiple hypothesis tests. We also investigate whether the presence of heterogeneity invalidates our estimated TEs.

4.1. Effect on random treatment group

To examine the effects of the training when trainees are randomly selected, we first estimate our models using the *first analysis sample*. Results are presented in Table 3 in which Panel (a) relies on our preferred district fixed effect model given by equation (1). Column 1 shows that training

[20] The method is double-robust, i.e., either the outcome or the treatment model can be misspecified but still can provide an unbiased estimate of the TE. Thus, the IPWRA estimates are valid even if our outcome model is wrong.

[21] AIPW estimators combine aspects of regression-adjustment and inverse-probability-weighted methods and have the double-robust property. Lasso, on the other hand, is a machine-learning approach to the selection of control variables. Selecting covariates with Lasso can be useful in two regards: to better deal with the power issues (Anderson & McKenzie, 2022) and to select a rich set of covariates and their interactions that can be correlated with treatment assignment (Bloniarz et al., 2016). As a result, the AIPW estimates obtained using Lasso are more likely to provide better estimates of the TE.

[22] Lee bound estimates an upper and a lower bound of the TEs by trimming, which corresponds to extreme assumptions about the missing values or the attrited observations.

[23] It can handle small samples and stratified treatment assignments, and thus indicate robustness of the results.

improves the probability of being gainfully employed by 18 percentage points (pp). The coefficient is large and statistically significant, indicating the success of the intervention in making the trainees employed.

[Table 3]

The results remain valid when we estimate the TE with RA, IPW, and AIPW with Lasso (panels b-d). The Lee bounds also confirm a significant effect on the outcome under the most conservative assumptions (panel e), while the randomization inference test results (panel f) confirm that the employed *t-distribution* based p-values are similar to those observed in our data.

Training does not seem to have any statistically and economically significant impact on the other considered in this analysis (Columns 2-6). For example, income only grows by 3.2 percent of the endline income of the RC group. Thus, our analysis indicates some impact of the training on employment but not on other outcomes. The lack of a significant increase in the total number of hours worked and income indicates that the benefit of being employed may not increase the the training participants' work length, and they can still be employed in low-paid jobs. The training has a limited (and likely insignificant) effect when we include home cultivation in defining gainful employment (Appendix Table A.5).[24]

The effect of vocational training only on employment is common in some previous studies (e.g., Barrera-Osorio et al., 2023). The pattern of findings can be explained by the fact that poor households in low-income countries are typically engaged in a portfolio of work rather than a single job (Blattman & Ralston, 2015). As a result, they may have the flexibility of reporting their employment status either way. So, we conclude that the effect of long training and incentive-based remuneration for the trainers results in a limited improvement in the outcomes and thus may not be very effective in improving the economic situation of the extreme poor in the short-run. These findings are consistent with most previous studies and reflect the fact that without capital,

[24]The table indicates a positive effect of training assignment on internal migration, indicating that improvement in employment can be through the domestic migration channel.

the returns to technical skills could be limited or that designing useful training programs can be challenging (Blattman & Ralston, 2015; McKenzie, 2017).[25,26]

4.2. Effect on trainer treatment group

Next, to examine the effects of the training when trainees are selected by the trainers, we estimate equation (1) using the *second analysis sample*. Results in Table 4 demonstrate the effects of training on all our chosen outcomes with our preferred model results presented in panel (a). The estimated TE in Column 1 indicates a 28 pp increase in the probability of gainful employment. The effect is about 50 percent higher than the impact on the RT group and is statistically significant at the 5 percent level against a one-sided alternative. The effect also remains significant when we employ RA (panel b), IPW (panel c), AIPW with Lasso (panel d), Lee bound (panel e), and randomization inference (panel f) in our analysis.

[Table 4]

Training also has a positive impact on the working hour of the TT group. The estimate indicates that the trainer-assigned trainees works 37 hours more per month than their counterparts (Column 2). The effect is also statistically significant at the 5 percent level against a one-sided alternative and is robust to the use of other methods employed earlier (panels b-f). Their income has also increased by NRs.2,550 (about 30 percent of the endline income of the RC group), although it is not statistically significant (Column 3). Understandably, their monthly income from working for themselves and business ownership is not affected (Columns 4-5) as they have been trained to get employed. However, their international migration significantly (against a one-sided alternative)

[25] Nevertheless, there are studies finding positive effects of vocational training in the short-run (e.g., Maitra & Mani, 2017; Doerr, 2022; Baird et al., 2022; Adhvaryu et al., 2023). Maitra & Mani (2017) find a subsidized vocational education program for women residing in low-income Indian households to increase participants' employment, working hour, and earnings in short- to medium-term. Doerr (2022) find that training vouchers in Germany translate into substantial gains in employment and earnings, specifically for low-skilled women. Baird et al. (2022) found an overall positive effect of randomized job training programs on earnings in New Orleans. Interestingly, some studies found an effect on the short-run that disappeared in the long-run (e.g., Hirshleifer et al., 2016; Blattman et al., 2020).

[26] The findings in Balboni et al. (2022) can be particularly useful in explaining the phenomenon. They examined whether people stay poor due to differences in fundamentals, such as ability, talent, or motivation, or differences in opportunities that stem from access to wealth. Using a large-scale, randomized asset transfer and an 11-year panel of 6,000 extreme poor households in rural Bangladesh, they find that above a threshold level of initial assets, households accumulate assets, take on better occupations, and grow out of poverty. but the reverse happens for those below the threshold.

increases by 7 pp compared to the RC group. The effect on international migration, however, fails to satisfy our employed robustness checks.

Thus, our analysis indicates some impact on the TT group on employment, working hours, and international migration. Robustness checks with similar outcome variables indicate a similar but less significant impact (Appendix Table A.6). So, we conclude that, in the short-run, trainers' selection generally provides a better outcome compared to the case when trainees are selected randomly. This can be due to the selected trainees' comparative advantage in vocational training, as observed in Silliman & Virtanen (2022).[27]

To better understand the finding and double-check the trainers' selection criteria, we later communicated with some training providers over the phone. They suggested that trainers would primarily look for the likelihood of applicants' taking a full-time job. They observed whether study participants' actions were consistent with their commitment to work. For instance, the trainers awarded higher scores to applicants visiting potential employers for job seeking. Similarly, trainers also favoured enthusiastic applicants who were even willing to pay the training fees, if required. Referrals from the previous cohort of trainees were also greatly valued. Some training providers gave priority to applicants who had family members already working in the same profession. This preference stem from the belief that familial connections could improve networking and, thus, increase prospects of employment. In short, they would try to delve deeper to gauge the attitude of the applicants.

4.3. Comparison of effects between treatment groups

At this state, we combine the data for the RC, RT and TT groups and estimate equation (2) to gain statistical significance from using a larger sample. Panel (a) results in Table 5 indicate a similar pattern of the impacts that we have observed earlier (Tables 3-4). Specifically, it provides a significant estimate of the effect on the employment of the RT group. The effects are again not statistically significant for any other outcome variables.In contrast, the TT group experiences

[27]Note that, for both RT and TT groups, we see some impact on employment but no effect on income, which appears puzzling. So, we investigate the issue further by looking at the impact on hours worked, wages earned, and monthly incomes of the groups, conditional on working, to understand the "intensive" margin. However, we have not found any meaningful significant impact in any of the cases. Results are available on request.

a significantly positive impact on employment and working hour. As expected, for both of the outcome variables, the estimated impacts are higher for the latter group.

[Table 5]

Next, we estimate model (3) to compare the differences in the impacts between the two treatment groups—RT and TT. Results in panel (b) of Table 5 indicate that the impact of training on the employment of the TT group is about 9 pp higher than the RT group. The difference is statistically significant at the 5 percent level against a one-sided alternative. The TT group also gains 15.4 working hours per month (again significant at the 5 percent level against a one-tailed test). Beneficial impacts on the other outcomes of interest are higher for the TT group but not statistically significant.[28]

Our estimated TEs are modest for the RT participants and are broadly consistent with Heckman et al. (1999) who suggest vocational training generate low benefits as they generally target low-quality participants. Our finding that the TT group experiences a (slightly) superior outcome is intuitive, as trainers may better understand applicants' ability, suitability, and motivation for jobs, as we have hypothesized. This is particularly so due to an incentive-based research design for the trainers. The pattern is broadly consistent with Rodríguez et al. (2022), who find the average returns to training vary across the unobserved ability distribution. The finding is also somewhat consistent with Campos et al. (2017), who, conducting an RCT in West Africa, find that personal initiative training, but not traditional training, improves outcomes.

4.4. Difference with the trainer control group and the effect of selection

To directly look into the effect of selection, we have also analyzed the changes in the outcomes for the TC group against the RC (reference) group. Members of both the groups do not take any training but the former one is negatively selected in some unobservable characteristics and may therefore experience deteriorated outcomes. When we use the *third analysis sample* and model (1) (now using the variable TC to indicate the trainer control group membership) for our purpose, we find no economically or statistically significant changes for the TC group (Appendix Table A.7).

[28] Results are similar when we include the TC group in our analysis.

Further robustness checks with competing outcome variables also find similar results (Appendix Table A.8).

Our previous analysis indicates that the TC group members generally do not experience a deteriorated outcome. This maybe because the applicants operate in the low-skilled job markets and, without any training, require only limited ability, motivation, and networking capacity. Thus, both the RC and TC group experience similar outcomes without vocational training.

Examining the direct effect of selection allows us to make a more useful interpretation of our earlier results. The effects of selection and training may not be additive, and so, when only better-quality applicants are trained, outcomes may improve through the following channels—a) the training, b) the quality (including better matching) of participants, and c) the interactions of training and quality. Our comparison of the RT group with the RC group (Table 3) offers an idea of (a). We also observe no significant effect on the TC group, applicants who were not selected by the trainers (Appendix Table A.7), indicating a likely limited contribution of (b). Thus, the improvement of the TT group over the RT group is likely due to the interaction effect (c). The overall results thus suggest that, while selection does not directly affect outcomes directly, it does so indirectly through the interaction with training. Such indirect effects may stem from, among others, the heterogeneous effect of training with regard to ability, matching, knowledge and network, which the trainers can guess during the trainee selection process.

A foolproof way to identify the causal effect of training on selected trainees can be achieved through other ways. For example, i) by randomly dividing the participants into treatment and control groups, and ii) for both of the groups, allowing the trainers to blindly (not knowing which is treatment and which is control group) select the trainees. The impact of training on a person the trainer would have chosen can then be estimated by comparing the two subgroups—trainer selected treatment and trainer selected control subgroups. While our setup is different, we have conducted some additional examination on this issue.

First, we compare the outcomes of the TT group with the part of the RC group whose trainer-provided scores were similar to the former group. The TE identified in this way is a substitute of the results from the ideal design, as the TT group members were selected solely on the scores, and so the two groups would have been similar except for training participation. When we compare the TT group with the part of the RC group whose members' scores are above the minimum score of the comparable (category-wise) TT group, we observe results that are similar to the pattern we observed earlier for the TT group (Appendix Table A.9). This indicates that, while selection is important to boost the TE, it alone does not contribute to improve the outcome of the training applicants, as we argue.

The impact of selection can be assessed in another way—by dividing the RC group into high-scoring and low-scoring subgroups. Comparing the two groups can indicate the effect of selection, as while all the applicants were selected randomly, those with high scores would have been selected by the trainers for training if they had been in the second part of the sample, retained to examine the selection issue; the opposite is true for those with low scores. We divide the RC group into two parts– one with members having category-wise above-mean scores, while the reference group includes the rest of them (with below-mean scores). We see null effects (against positive alternatives) of positive selection, indicating no difference in outcomes between the two groups (Appendix Table A.10). This reassures us that selection does not directly contribute to improving the outcome of the training participants.

To better compare the size of the effects on the outcomes between RT and TT groups, Figure 1 below presents the standardized effects/changes on all six outcomes, summarising our earlier results.

[Figure 1]

It is worth discussing a potential implication of our research design on the estimates of the TEs. To fully utilize the available capacity, the trainers assign around 72 percent of the participants to the training. The mechanism is likely to be less successful in selecting better quality participants than a case, for example, that selects (top) 10 percent of the participants for training. Thus, by design, our experiment is likely to find a lower effect of training against more conservative selections.

4.5. Robustness and Heterogeneity

At this stage, we conduct some additional examination to confirm that our estimated effects are robust. First, our investigation relied on multiple outcomes of interest, which raises the issue of false discovery rate (FDR) associated with multiple hypothesis testing (List et al., 2019, 2023). To report the correct significance level (adjusted for multiple hypothesis testing) of our exposure variable, we follow the approach provided in Romano & Wolf (2005a,b, 2016); Clarke et al. (2020).

Table 6 reports three types of *p-values* for the TEs (or changes) in the outcome variables for each of the RT, TT and TC groups. Columns 1 and 4 present simple (uncorrelated model) *p-values*, Columns 2 and 5 present the *p-values* by random permutation respecting strata and clusters, while Columns 3 and 6 present the Romano-Wolf (R-W) multiple hypotheses corrected *p-values*.

[Table 6]

Our conclusions remain unaffected with the use of any, including R-W multiple hypotheses corrected *p-values*, indicating statistically significant effects of the training on i) employment for the RT group, and ii) employment and working hour on the TT group. The table also presents randomization *p-values* for joint tests of treatment significance, as discussed in Young (2019). As earlier, for both groups we reject the null hypothesis that training improves none of the outcomes, confirming positive effects of the training.

Next, we confirm that the exposure variables are not made significant by *p-hacking*. To do so, we use the method in Brodeur et al. (2020a,b) and check whether the use of various combinations of control variables changes the significance of the coefficient of the treatment/group dummy variables. We generate standardized graphical outputs from regression specifications by individually regressing a dependent variable against all possible combinations of independent variables (Appendix Figures A.1-A.2). The effect curves (histograms of the estimated TEs) and the *t-curves* (histograms of the absolute value of the *t-statistics* of the TEs) closely match with the estimates from our employed models, suggesting towards the validity of our estimates.[29]

McKenzie (2017) suggest that the real impact of vocational training is small and thus difficult to identify when the sample sizes are small. Our study employs a reasonable training duration and the sample sizes also appear reasonable.[30] Nevertheless, for each outcome, we examine whether our study design have enough power to detect a modest effect. To do so, based on observed standard deviations in the actual outcome of the RC groups, we compute the minimum detectable effect size (MDES) with adequate statistical power. We follow the standard practice of 80 percent power with

[29]See appendix, Figures A.1-A.2, where we presented the standardized graphical output for all the six outcome variables (in the same order, from left to right and top to bottom) for the RT and TT groups. The specification tests used the Stata code "*speccheck*" provided by the authors of Brodeur et al. (2020a,b) in https://sites.google.com/site/abelbrodeur/speccheck.

[30]Agarwal & Mani (2024) found that the duration of recent experimental studies on vocational training program ranges from 1-48 months, depending on the type of the program.

a two-sided test at 5 percent significance level (Islam et al., 2021). A true positive impact smaller than the corresponding MDES will have less than 80 percent chance of being identified.

For the RT group, our estimated effect size is larger than MDES (in their original units of measurement) only for employment (Appendix Table A.11). For the TT group, the estimated effects are larger than MDES for employment and working hour. Thus, we may fail to detect the positive effect of training on some outcomes of interest. However, the comparison of MDES with our estimated effects supports our conclusion of a lower effect of vocational training on the RT group compared to the TT group.

Noncompliance is always an important issue in interpreting the results derived from RCT designs. In our case, some RC group members took the training, while the opposite is true for some RT and TT group members, raising the the issue that the ITT estimates are likely to underestimate the true TEs. Therefore, using group wise training assignment as an instrument of actual (group-wise) training participation, we estimate the LATEs of training participation on all the outcomes. Our estimation of LATE otherwise follows specification (3). As expected, the LATE estimates are higher than their ITT estimates for both the RT and TT groups. Also, in all cases, the impact on the TT group is higher than the RT group, although the differences are statistically significant only for employment and working hour (Appendix Table A.12). Overall, our conclusions about the effectiveness of training for the trainer-selected group remain unaffected when we consider the statistical significance of the estimated TEs.

Heterogeneity in the TEs is commonly observed in empirical studies on vocational training (Blattman & Ralston, 2015; McKenzie, 2017; Mckenzie, 2023). Specifically, average returns to training vary across sex (Acevedo et al., 2020; Attanasio et al., 2011), education (Kiuma et al., 2020), income and wealth (Galdo et al., 2008), caste (Field et al., 2010) and unobserved ability (Rodríguez et al., 2022). One particular problem is that the OLS estimation of equation (1) is generally inappropriate in the presence of heterogeneity (Słoczyński, 2022). To examine whether heterogeneity is a threat to our estimated TEs (or changes for the case of TC group), we repeat the previous analysis by sex, education, and income subgroups.[31]

[31] We could neither investigate the treatment heterogeneity by caste due to a small subsample size nor by unobserved heterogeneity due to data unavailability. The effect of vocational training also depends critically on program design and delivery elements (Carranza & Mckenzie, 2023).

Table 7 shows the estimated TEs (and their SEs) for the outcome variables by analysis samples and subgroups defined by sex, education, and income (details in Appendix Tables B.1-B.6). Ignoring the statistical significance for now, we usually observe a positive impact on employment for all subgroups of the RT group (panel a). Moreover, female, low-educated, and low-income participants benefit more. On the other hand, we generally observe a positive impact on employment, working hour, and income for all subgroups of the TT group (panel b). Again, female, low-educated, and low-income participants benefit more. For all subgroups, the effects are mostly higher for the TT group compared to the RT group.

[Table 7]

The higher impact on females is consistent with Attanasio et al. (2011), who find vocational training raises earnings and employment for women in Colombia, and with Acevedo et al. (2020) who find strong and lasting effects of soft skills training on personal skills acquisition and expectations for women but not for men. The pattern of differential impact between men and women suggests that the success of job-training programs may depend on trainees' expectations, as found in Acevedo et al. (2020).[32] Education is also likely to interact with the training positively through productivity and negatively through motivation. For example, Bassanini (2004) find training has a stronger impact on employment security for low-educated workers.

The higher effect on low-income individuals may be due to their motivation and urgency in finding jobs to survive. For example, Doerr (2022) find that low-skilled workers benefited most from a vocational training program in Germany. The same is true for low-income trainees, as they are likely to be low-skilled. While our estimated effects are largely statistically insignificant they still demonstrate a generally larger impact on the TT group compared to the RT group.[33]

We also estimate models with interactions of group dummies and trainee characteristics and conclude similarly (Appendix Table B.7). To further confirm that heterogeneity does not invalidate the estimated TEs, we use the method provided in Słoczyński (2022). The results largely indicate

[32] A randomised experiment in India found that including information sessions about placement opportunities make vocational trainees more likely to stay in the jobs in which they are placed, as trainees who are over-optimistic about placement jobs are more likely to drop out before placement (Chakravorty et al., 2024).

[33] Note that discovering and exploiting TE heterogeneity is not a goal of this research. The subgroup analysis here aims to show whether our results are robust even after considering subgroup heterogeneity and whether they can shed additional light as suggested in Duflo et al. (2007). Discovering and exploiting heterogeneity of the TEs requires ex-ante specification of and random assignment into subgroups, requiring a larger sample than that required for examining whether the treatment has an effect (Duflo et al., 2007; Chernozhukov et al., 2018; List, 2025).

that the estimated TEs are similar to the estimated ATT or ATE (Appendix Table B.8). Thus, the heterogeneity analysis is consistent with the literature and supports our findings— vocational training may provide extra benefits in the short run when trainers select their participants.

5. Policy relevance and intervention cost recovery

Developing countries around the world are continuously seeking ways to improve the economic status of their populations at the bottom of the income distribution. In this regard, vocational training, which we evaluate here, is an approach to enhance labor productivity and thereby increase their employment opportunities. Thus, our findings may have important policy implications in this context. Firstly, we confirm that even intensive vocational training, combined with trainers' incentives linked to trainees' employment, can only affect employment prospect in the short-run. This finding is consistent with a large number of studies reporting a null or small effects of such training (e.g., Heckman et al., 1999; Blattman & Ralston, 2015; McKenzie, 2017).[34,35]

We also find that the impact on employment is higher, while the working hour increases when the trainers select trainees. The LATE estimates indicate that actual TEs are higher than our estimated ITT effects. The reason for elevated impact on the trainer selected group can be due to the trainers' local knowledge of unobservable trainee characteristics, such as drive, networks, matching ability, and motivation for jobs and income. Together with the incentive-based payment, these unobservable characteristics can effectively improve outcomes like working hour and income. The improved outcomes may also be due to signalling and realization of quality, as previous studies have found that individuals possess valuable skills unobservable to employers (e.g., Abebe et al., 2021a,b).

Therefore, our study has significant implications for the design of training programs, suggesting that trainers should select the training participants, particularly when their remuneration is tied to the trainees' job market performance. In doing so, we contribute to the targeting literature, which parallels the causal machine learning literature evaluating TE heterogeneity (e.g., Aiken

[34] Vocational training may have some other beneficial effects on society. For example, skill development training programs for women contribute to liberalizing the gender norms and attitudes around women working outside the household (Janzen et al., 2021). While those objectives are vital, this study focuses solely on economic outcomes.

[35] A recent meta analysis by Agarwal & Mani (2024) finds a small overall effect of vocational training, although they estimate a null effect for vocational plus on the job training—the type of training considered in our study.

et al., 2022, 2023; Athey et al., 2023), though we focus on the trainee characteristics that are either unobservable (without an interview) or difficult to measure.

A proper cost-benefit analysis framework, however, compares the program cost against the estimated benefits of the training. The estimated benefit of the training on monthly income is NRs.270 for the randomly selected trainees and NRs.2,550 for the trainer-selected applicants. Although none of the estimates are statistically significant, the latter one is economically large. With a training cost of NRs.40,000 per trainee, our back-of-the-envelope calculation indicates that the former group requires 12 years, while the latter group needs about one year and a quarter to recover the training cost. The effect on the income of trainer selected group is notably larger than in the most recent RCT-based vocational training studies listed in Agarwal & Mani (2024).[36]

We can also take the return-on-investment approach discussed in detail in McKenzie (2021). With a five percent monthly return from investing in a microenterprise, as suggested by De Mel et al. (2008), financing the training cost would earn NRs.2,000 per month. This return appears to be much higher than the increase in income of the randomly selected trainees, but the opposite is true for the trainer-selected applicants. Even a one percent monthly return provides a higher benefit than the gain in income for the randomly selected trainees. The cost-benefit analysis thus indicates that vocational training on the randomly selected trainees misallocates resources, which is counterproductive. In contrast, the training of the trainer-selected applicants is productive.

One key concern with these types of job training programs is that they may dispace rather than create new jobs (McKenzie, 2017; Mckenzie, 2023). This pattern has been observed in some previous studies like Crépon et al. (2013). Howevr, the possibility of crowding out seems less likely in our case, as we have observed in the trainer-selected trainee group that the training raises international migration. Previous studies have found large benefits from out-migration, including benefits to the people in the location of origin (Bryan et al., 2014; Meghir et al., 2022). Thus, it is likely that by inducing out-migration, the training increases participants' benefit without negatively affecting others already working in that field.

[36] Incorporating the domestic interest rate into the analysis, which is also more appropriate, would further increase the time required to recover the training cost, but for simplicity, we have excluded this factor from our calculations.

6. Conclusions

We investigate the short-run impact of intensive vocational training with trainers' incentives on applicants' employment, working hours, income, business ownership, and international migration. We find that the training has limited effects on these outcomes, but can generate some benefits when the trainees are selected by the trainers, who may have some insight into unobservable characteristics such as motivation, knowledge, matching and ability of the trainees. Our results suggest that positive selection did not directly affect the outcomes but influenced them indirectly through the interaction with the training. Trainers, during the interview, can gauge applicant's qualities that enhance the training outcomes, leading to positive selection. Interviews with the trainers later confirmed our understanding of why the training had a greater impact on selected applicants.

The cost-benefit analysis indicates that the randomly selected vocational training participants require a long time to recoup their training cost, while the trainer selected trainees can recover costs much more quickly. As this study is one of the most rigorous evaluations of vocational training with an RCT research design in Nepal, its credible findings can assist Nepal and other low-income countries in designing policies to promote employment and reduce poverty. It may also attract the interest of key stakeholders, including training providers, NGOs, government agencies, and international donors.

Thus, it is worth discussing the potential for scaling up the training programs. We follow the five criteria suggested by List (2022) for scaling. The first criterion is whether more evidence is needed before scaling. For the trainer-selected participants, since we chose a large proportion of applicants for training, it is worth exploring whether outcomes could improve with a smaller proportion of training participants. Thus, more evidence is needed in this regard before scaling up. The second criterion is whether the samples used are representative of the population. This criterion is satisfied, as we randomly selected applicants from the interested participants across Nepal. The sample is likely similar to those in other developing countries with comparable settings.

The third criterion is whether the intervention conducted under conditions representative of the broader situation. So far, the conditions in Nepaland other developing countries, particularly in South Asia, are similar, where there are large pools of applicants interested in traditional jobs.

Thus, the vocational training from our study could benefit many low-income youths. The fourth criterion is whether there are likely spillover (network) effects and general equilibrium (GE) effects from scaling up. We observed no effect of the training on other family members' income, suggesting that spillover effects are unlikely to be negative in the long run as long, provided that the trained participants do not replace their untrained competitors in the job market—an issue we discussed earlier. Scaling up may, however, bring a positive GE effect by reducing poverty and vulnerability in the region. The final criterion is whether any diseconomies of scale are associated with the intervention. Since our training contents and mechanisms are simple, it is easy to train additional trainers, implying that the intervention is likely to avoid diseconomies of scale.

Nonetheless, our approach to estimate the causal effect of trainee selection could be improved further by randomly choosing treatment and control participants in the first stage, then allowing trainers to select the trainees without knowing their groups in the second stage. The impact of training then could be more credibly estimated by comparing the two trainer-selected subgroups – one from the treatment and another from the control group. Therefore, we recommend more investigation into the impact of vocational training using this design mentioned above before strongly recommending scaling up the program. This will also allow for "Option C Thinking", as suggested by List (2024).

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1), 1–35.

Abebe, G., Caria, A. S., Fafchamps, M., Falco, P., Franklin, S., & Quinn, S. (2021a). Anonymity or distance? Job search and labour market exclusion in a growing African city. *Review of Economic Studies*, 88(3), 1279–1310.

Abebe, G., Caria, A. S., & Ortiz-Ospina, E. (2021b). Selection of talent: Experimental and structural evidence from Ethiopia. *American Economic Review*, 111(6), 1757–1806.

Acevedo, P., Cruces, G., Gertler, P., & Martinez, S. (2020). How job training made women better off and men worse off. *Labour Economics*, 65(2020), 101824.

Adhvaryu, A., Kala, N., & Nyshadham, A. (2023). Returns to on-the-job soft skills training. *Journal of Political Economy*, 131(8), 2165–2208.

Agarwal, N. & Mani, S. (2024). New evidence on vocational and apprenticeship training programs in developing countries. *Handbook of Experimental Development Economics*, Forthcoming. Web: https://dx.doi.org/10.2139/ssrn.4851428 [Viewed at: 4 August, 2024].

Aiken, E., Bellue, S., Karlan, D., Udry, C., & Blumenstock, J. E. (2022). Machine learning and phone data can improve targeting of humanitarian aid. *Nature*, 603(7903), 864–870.

Aiken, E. L., Bedoya, G., Blumenstock, J. E., & Coville, A. (2023). Program targeting with machine learning and mobile phone data: Evidence from an anti-poverty intervention in Afghanistan. *Journal of Development Economics*, 161, 103016.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., & Vitali, A. (2020). Tackling youth unemployment: Evidence from a labor market experiment in Uganda. *Econometrica*, 88(6), 2369–2414.

Allcott, H. (2015). Site selection bias in program evaluation. *Quarterly Journal of Economics*, 130(3), 1117–1165.

Anderson, S. J. & McKenzie, D. (2022). Improving business practices and the boundary of the entrepreneur: A randomized experiment comparing training, consulting, insourcing, and outsourcing. *Journal of Political Economy*, 130(1), 157–209.

Asian Development Bank (2017). *Country Poverty Analysis (Detailed): Nepal.* Asian Development Bank (ADB). Web: https://www.adb.org/countries/nepal/overview, [Viewed at: 18 April, 2022].

Athey, S. & Imbens, G. W. (2017). The econometrics of randomized experiments. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of Economic Field Experiments*, volume 1 (pp. 73–140). North-Holland.

Athey, S., Keleher, N., & Spiess, J. (2023). Machine learning who to nudge: Causal vs predictive targeting in a field experiment on student financial aid renewal. *arXiv preprint arXiv:2310.08672*. Web: https://arxiv.org/pdf/2310.08672.pdf, [Viewed at: 22 October, 2023].

Attanasio, O., Guarín, A., Medina, C., & Meghir, C. (2017). Vocational training for disadvantaged youth in Colombia: A long-term follow-up. *American Economic Journal: Applied Economics*, 9(2), 131–43.

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

Baird, M. D., Engberg, J., & Gutierrez, I. A. (2022). RCT evidence on differential impact of US job training programmes by pre-training employment status. *Labour Economics*, 75, 102140.

Balboni, C., Bandiera, O., Burgess, R., Ghatak, M., & Heil, A. (2022). Why do people stay poor? *Quarterly Journal of Economics*, 137(2), 785–844.

Banerjee, A., Duflo, E., & Hornbeck, R. (2018). How much do existing borrowers value microfinance? Evidence from an experiment on bundling microcredit and insurance. *Economica*, 85(340), 671–700.

Banerjee, A., Karlan, D., & Zinman, J. (2015). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7(1), 1–21.

Barrera-Osorio, F., Kugler, A., & Silliman, M. (2023). Hard and soft skills in vocational training: Experimental evidence from Colombia. *World Bank Economic Review*, 37(3), 409–436.

Bassanini, A. (2004). Improving skills for more and better jobs? *European Economy: Special Reports*, 3(8), 103–137.

Blattman, C., Fiala, N., & Martinez, S. (2013). Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *Quarterly Journal of Economics*, 129(2), 697–752.

Blattman, C., Fiala, N., & Martinez, S. (2020). The long-term impacts of grants on poverty: Nine-year evidence from Uganda's youth opportunities program. *American Economic Review: Insights*, 2(3), 287–304.

Blattman, C. & Ralston, L. (2015). *Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs*. White paper, Prepared for the World Bank, Washington DC, USA.

Bloniarz, A., Liu, H., Zhang, C.-H., Sekhon, J. S., & Yu, B. (2016). Lasso adjustments of treatment effect estimates in randomized experiments. *Proceedings of the National Academy of Sciences*, 113(27), 7383–7390.

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264), 1073–1076.

Brodeur, A., Cook, N., & Heyes, A. (2020a). A proposed specification check for p-hacking. *AEA Papers and Proceedings*, 110, 66–69.

Brodeur, A., Cook, N., & Heyes, A. (2020b). Methods matter: P-hacking and publication bias in causal analysis in economics. *American Economic Review*, 110(11), 3634–60.

Bryan, G., Chowdhury, S., & Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5), 1671–1748.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., & Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small business in West Africa. *Science*, 357(6357), 1287–1290.

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

Card, D., Kluve, J., & Weber, A. (2018). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3), 894–931.

Carranza, E. & Mckenzie, D. J. (2023). *Job Training and Job Search Assistance Policies in Developing Countries*. IZA Discussion Paper 16537, IZA – Institute of Labor Economics, Bonn, Germany.

Chakravarty, S., Lundberg, M., Nikolov, P., & Zenker, J. (2019). Vocational training programs and youth labor market outcomes: Evidence from Nepal. *Journal of Development Economics*, 136, 71–110.

Chakravorty, B., Arulampalam, W., Bhatiya, A. Y., Imbert, C., & Rathelot, R. (2024). Can information about jobs improve the effectiveness of vocational training? Experimental evidence from India. *Journal of Development Economics*, 169, 103273.

Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I. (2018). *Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments*. NBER Working Paper 24678, National Bureau of Economic Research, Cambridge, MA, USA.

Cho, Y. & Honorati, M. (2014). Entrepreneurship programs in developing countries: A meta regression analysis. *Labour Economics*, 28, 110–130.

Cho, Y., Kalomba, D., Mobarak, A. M., & Orozco-Olvera, V. (2013). *Gender differences in the effects of vocational training: Constraints on women and drop-out behavior*. Policy Research Working Paper 6545, World Bank, Washington DC, USA.

Clarke, D., Romano, J. P., & Wolf, M. (2020). The Romano–Wolf multiple-hypothesis correction in Stata. *Stata Journal*, 20(4), 812–843.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., & Zamora, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Quarterly Journal of Economics*, 128(2), 531–580.

Das, N. (2021). Training the disadvantaged youth and labor market outcomes: Evidence from Bangladesh. *Journal of Development Economics*, 149, 102585.

De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *Quarterly Journal of Economics*, 123(4), 1329–1372.

Doerr, A. (2022). Vocational training for female job returners - effects on employment, earnings and job quality. *Labour Economics*, 75, 102139.

Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. In T. P. Schultz & J. A. Strauss (Eds.), *Handbook of Development Economics*, volume 4 (pp. 3895–3962). Elsevier.

Employment Fund (2013). *Path to Prosperity*. Implementation guideline, Employment Fund Secretariat, HELVETAS Swiss Intercooperation, Kathmandu, Nepal.

Field, E., Jayachandran, S., & Pande, R. (2010). Do traditional institutions constrain female entrepreneurship? A field experiment on business training in India. *American Economic Review*, 100(2), 125–29.

Galdo, J., Jaramillo, M., & Montalva, V. S. (2008). Household wealth and heterogeneous impacts of a market-based training program: the Case of PROJOVEN in Peru. *Economics Faculty Scholarship*, 143. Web: http://surface.syr.edu/ecn/143 [Viewed at: 15 August, 2024].

Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics*, volume 3 (pp. 1865–2097). Elsevier.

Heß, S. (2017). Randomization inference with Stata: A guide and software. *Stata Journal*, 17(3), 630–651.

Hirshleifer, S., McKenzie, D., Almeida, R., & Ridao-Cano, C. (2016). The impact of vocational training for the unemployed: Experimental evidence from Turkey. *Economic Journal*, 126(597), 2115–2146.

Islam, A., Lee, W.-S., & Nicholas, A. (2021). The effects of chess instruction on academic and non-cognitive outcomes: Field experimental evidence from a developing country. *Journal of Development Economics*, 150, 102615.

J-PAL (2022). *Data analysis.* The Abdul Latif Jameel Poverty Action Lab (J-PAL). Web: https://www.povertyactionlab.org/resource/data-analysis, [Viewed at: 18 April, 2022].

Janzen, S. A., Magnan, N., Mullally, C. C., & Sharma, S. (2021). *Training and shifting gender norms: Evidence from a training intervention in rural Nepal.* Paper presented at the AAEA Annual Meeting, August 1-3, Agricultural & Applied Economics Association, Austin, TX, USA.

Katz, L. F., Roth, J., Hendra, R., & Schaberg, K. (2022). Why do sectoral employment programs work? Lessons from WorkAdvance. *Journal of Labor Economics*, 40(S1), S249–S291.

Khandker, S. R. (2005). Microfinance and poverty: Evidence using panel data from Bangladesh. *World Bank Economic Review*, 19(2), 263–286.

Kiuma, A. K., Araar, A., & Kaghoma, C. K. (2020). Internal migration and youth entrepreneurship in the Democratic Republic of the Congo. *Review of Development Economics*, 24(3), 790–814.

Krauss, A. (2018). Why all randomised controlled trials produce biased results. *Annals of Medicine*, 50(4), 312–322.

Krauss, A. (2021). Assessing the overall validity of randomised controlled trials. *International Studies in the Philosophy of Science*, 34(3), 159–182.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3), 1071–1102.

List, J. A. (2022). *The Voltage Effect.* Penguin, UK, 1st edition.

List, J. A. (2024). Optimally generate policy-based evidence before scaling. *Nature*, 626(7999), 491–499.

List, J. A. (2025). Field experiments: Here today gone tomorrow? *The American Economist*, Forthcoming, 05694345241261340.

List, J. A., Shaikh, A. M., & Vayalinkal, A. (2023). Multiple testing with covariate adjustment in experimental economics. *Journal of Applied Econometrics*, 38(6), 920–939.

List, J. A., Shaikh, A. M., & Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4), 773–793.

Maitra, P. & Mani, S. (2017). Learning and earning: Evidence from a randomized evaluation in India. *Labour Economics*, 45, 116–130.

McKenzie, D. (2017). How effective are active labor market policies in developing countries? A critical review of recent evidence. *World Bank Research Observer*, 32(2), 127–154.

McKenzie, D. (2021). Small business training to improve management practices in developing countries: Re-assessing the evidence for 'training doesn't work'. *Oxford Review of Economic Policy*, 37(2), 276–301.

Mckenzie, D. J. (2023). *Is There Still A Role for Direct Government Support to Firms in Developing Countries?* Policy Research Working Paper 10628, World Bank, Washingto DC, USA.

Meghir, C., Mobarak, A. M., Mommaerts, C., & Morten, M. (2022). Migration and informal insurance: Evidence from a randomized controlled trial and a structural model. *Review of Economic Studies*, 89(1), 452–480.

Pitt, M. M. & Khandker, S. R. (1998). The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter? *Journal of Political Economy*, 106(5), 958–996.

Rodríguez, J., Saltiel, F., & Urzúa, S. S. (2022). Dynamic treatment effects of job training. *Journal of Applied Econometrics*, 37(2), 242–269.

Romano, J. P. & Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469), 94–108.

Romano, J. P. & Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4), 1237–1282.

Romano, J. P. & Wolf, M. (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, 113, 38–40.

Silliman, M. & Virtanen, H. (2022). Labor market returns to vocational secondary education. *American Economic Journal: Applied Economics*, 14(1), 197–224.

Słoczyński, T. (2022). Interpreting OLS estimands when treatment effects are heterogeneous: Smaller groups get larger weights. *Review of Economics and Statistics*, 104(3), 501–509.

Van den Berg, G. J. & Vikström, J. (2022). Long-run effects of dynamically assigned treatments: A new methodology and an evaluation of training effects on earnings. *Econometrica*, 90(3), 1337–1354.

Wooldridge, J. M. (2021). *Introductory econometrics: A modern approach.* Cengage learning, 2$^{\text{nd}}$ edition.

Young, A. (2019). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *Quarterly Journal of Economics*, 134(2), 557–598.

Tables and Figures

TABLE 1: **Selected applicants by assignment type**

Group type	Observations
a. Random control (RC)	139
b. Random treatment (RT)	373
c. Trainer treatment (TT)	374
d. Trainer control (TC)	150
Total program participants	1,036

Note: (a+b) makes our *first analysis sample*; (a+c) makes our *second analysis sample*; and (a+d) makes our *third analysis sample*. Total program participants is given by (a+b+c+d).

TABLE 2: **Balance during baseline by group type**

Variable \ Group	Random control	Random treatment		Trainer treatment		Trainer control	
	Mean (1)	Mean (2)	Difference (3)	Mean (4)	Difference (5)	Mean (6)	Difference (7)
a. Control variables							
Age 15-19	0.23	0.20	-0.03	0.22	-0.01	0.19	-0.04
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.03)	(0.05)
Age 20-24	0.31	0.35	0.04	0.32	0.01	0.41	0.10*
	(0.04)	(0.02)	(0.05)	(0.02)	(0.05)	(0.04)	(0.06)
Age 25-29	0.23	0.19	-0.04	0.21	-0.02	0.19	-0.04
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.03)	(0.05)
Age 30-34	0.14	0.17	0.02	0.14	0.00	0.14	-0.00
	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)	(0.04)
Age 35-39	0.08	0.09	0.01	0.09	0.01	0.05	-0.03
	(0.02)	(0.01)	(0.03)	(0.01)	(0.03)	(0.02)	(0.03)
Age 40-49	0.01	0.01	0.00	0.01	0.00	0.02	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female	0.34	0.36	0.02	0.41	0.07	0.25	-0.08
	(0.04)	(0.02)	(0.05)	(0.03)	(0.05)	(0.04)	(0.05)
Education: below primary	0.30	0.25	-0.05	0.24	-0.06	0.22	-0.08
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.03)	(0.05)
Education: primary to below SLC	0.42	0.46	0.04	0.51	0.10*	0.51	0.09
	(0.04)	(0.03)	(0.05)	(0.03)	(0.05)	(0.04)	(0.06)
Education: SLC and beyond	0.28	0.29	0.01	0.24	-0.04	0.27	-0.01
	(0.04)	(0.02)	(0.05)	(0.02)	(0.04)	(0.04)	(0.05)
Never married	0.41	0.38	-0.03	0.34	-0.07	0.44	0.03
	(0.04)	(0.03)	(0.05)	(0.02)	(0.05)	(0.04)	(0.06)
Brahmin and Chhetri	0.19	0.24	0.05	0.20	0.02	0.21	0.02
	(0.03)	(0.02)	(0.04)	(0.02)	(0.04)	(0.03)	(0.05)
Prior training participation	0.08	0.06	-0.02	0.09	0.01	0.09	0.01
	(0.02)	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)
F-test (p-value)	-	-	0.76	-	0.33	-	0.64
Observations	139	373	512	374	513	150	289
b. Outcome variables							
Gainfully employed	0.36	0.29	-0.06	0.26	-0.10**	0.33	-0.03
	(0.04)	(0.02)	(0.05)	(0.02)	(0.04)	(0.04)	(0.06)
Monthly hours worked	117.22	119.38	2.16	119.59	2.37	104.13	-13.09
	(8.26)	(5.08)	(9.73)	(5.69)	(10.61)	(9.21)	(12.44)
Monthly own income	3.43	2.54	-0.90	2.21	-1.22**	1.77	-1.67**
	(0.72)	(0.28)	(0.64)	(0.26)	(0.61)	(0.25)	(0.74)
Income working for oneself	1.87	1.16	-0.71	0.99	-0.87	0.65	-1.22*
	(0.66)	(0.25)	(0.58)	(0.22)	(0.54)	(0.19)	(0.67)
Owns business	0.11	0.11	0.00	0.09	-0.02	0.08	-0.03
	(0.03)	(0.02)	(0.03)	(0.01)	(0.03)	(0.02)	(0.03)
International Migration	0.02	0.04	0.02	0.02	0.00	0.05	0.03
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
F-test (p-value)	-	-	0.57	-	0.27	-	0.38
Observations	139	373	512	374	513	150	289

Note: Means are reported; SEs are in the parentheses. Column 3 shows the difference between RT and the RC group; column 5 shows the same between TT and RC group and column 7 shows the same between TC and RC group. The *s indicate the p-values from the t-tests of differences in the means across the groups (against a two-sided alternative): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The F-test of joint significance runs a regression of the relevant group dummy on all the outcome variables and then tests the null hypothesis that all the slope coefficients are zero. The specific control variables related to the applicants are, age in years (that we categorized as 15-19, 20-24, 25-29, 30-34, 35-39, and 40-49 years), whether female, years of education (that we categorized as below primary, primary to below SLC and SLC or beyond), whether married and whether belong to upper caste (Brahmin or Chhetri). School Leaving Certificates (SLCs) are given after completing Grade 10. For more details about the education system in Nepal, see https://www.scholaro.com/pro/Countries/Nepal/Education-System. We also control for prior participation in vocational or skill training. Monetary variables are in thousand Nepalese Rupees. The definition of the variable "Gainfully employed" excludes home cultivation, a proxy for subsistence farming.

Table 3: **ITT effect on random treatment group**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. With district fixed effects						
Treatment	0.18**	14.42	0.27	-0.58	-0.00	0.00
	(0.08)	(10.84)	(1.43)	(0.90)	(0.05)	(0.04)
b. With regression adjustment						
Treatment	0.19***	13.80	1.02	0.06	-0.02	0.02
	(0.05)	(9.56)	(1.50)	(1.31)	(0.04)	(0.04)
c. With inverse probability weighting						
Treatment	0.19***	13.80	1.02	0.06	-0.02	0.02
	(0.05)	(9.56)	(1.50)	(1.31)	(0.04)	(0.04)
d. With selection of covariates using Lasso						
Treatment	0.17***	14.36	1.02	0.16	0.01	0.02
	(0.05)	(9.79)	(1.73)	(1.54)	(0.04)	(0.04)
e. Lee bounds						
lower	0.14***	-3.60	-2.07	-2.07	-0.06	-0.08
	(0.05)	(11.88)	(2.26)	(2.12)	(0.04)	(0.06)
upper	0.27***	32.31***	2.03	0.46	0.02	0.05
	(0.05)	(12.11)	(1.90)	(1.62)	(0.04)	(0.04)
f. Significance level with randomization inference						
Treatment	0.18***	14.42	0.27	-0.58	-0.00	0.00
g. Random control mean						
At endline	0.61	171.07	8.46	3.07	0.14	0.15
N	461	419	442	442	442	461

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34, 35-39, and 40-49 years), gender, education (below primary, primary and secondary), marital status, caste, prior training experience and the value of the outcome variable at the baseline as well as the district fixed effects. SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * p <0.10, ** p <0.05, *** p <0.01. We used Stata command *"teffects ra"* to estimates the regression adjusted TEs, *"teffects ipwra"* to estimate inverse probability weighted regression adjusted TEs and *"telasso"* to estimate inverse-probability weighted TEs that also use the LASSO method to select the control variables to be included in the model. We use Stata command *"leebounds"* to estimates the Lee bounds of the TEs as suggested by Lee (2009). We used an unofficial Stata command *"ritest"* to estimate the randomization inference significance levels and p-values. The command is written by Heß (2017) that is freely available from https://github.com/simonheb/ritest.

Table 4: **ITT effect on trainer treatment group**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. With district fixed effects						
Treatment	0.28*	36.93*	2.55	-0.24	0.01	0.07*
	(0.14)	(18.98)	(2.37)	(1.07)	(0.09)	(0.04)
b. With regression adjustment						
Treatment	0.22***	17.41*	1.53	0.21	0.07*	0.01
	(0.05)	(9.85)	(1.59)	(1.15)	(0.04)	(0.04)
c. With inverse probability weighting						
Treatment	0.22***	17.41*	1.53	0.21	0.07*	0.01
	(0.05)	(9.85)	(1.59)	(1.15)	(0.04)	(0.04)
d. With selection of covariates using Lasso						
Treatment	0.19***	17.25*	1.22	0.18	0.09**	0.00
	(0.05)	(9.92)	(1.74)	(1.36)	(0.04)	(0.04)
e. Lee bounds						
lower	0.18***	7.65	-0.67	-0.60	0.04	-0.07
	(0.05)	(12.81)	(2.15)	(1.82)	(0.06)	(0.06)
upper	0.28***	25.53**	2.23	0.39	0.09**	0.04
	(0.07)	(12.71)	(2.01)	(1.45)	(0.04)	(0.04)
f. With randomization inference						
Treatment	0.28***	36.93***	2.55	-0.24	0.01	0.07*
g. Random control mean						
At endline	0.61	171.07	8.46	3.07	0.14	0.15
N	453	404	429	429	429	453

Note: See the notes in Table 3.

Table 5: **ITT effect of training by group type**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. With separate dummies for the treatment groups						
Random treatment	0.17**	13.58	-0.29	-0.89	-0.01	0.01
group	(0.08)	(10.19)	(1.23)	(0.85)	(0.05)	(0.04)
Trainer treatment	0.26**	28.82**	3.57	1.54	0.05	0.05
group	(0.09)	(13.71)	(2.97)	(2.06)	(0.09)	(0.06)
b. With treatment dummy and its' interaction with the TT group dummy						
Treatment	0.17**	13.58	-0.29	-0.89	-0.01	0.01
	(0.08)	(10.19)	(1.23)	(0.85)	(0.05)	(0.04)
Treatment $\times$ TT	0.09*	15.25*	3.87	2.43	0.06	0.04
	(0.05)	(8.80)	(2.96)	(2.45)	(0.07)	(0.08)
N	800	720	759	759	759	800

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34, 35-39, and 40-49 years), gender, education (below primary, primary and secondary), marital status, caste, prior training experience and the value of the outcome variable at the baseline as well as the district fixed effects. SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

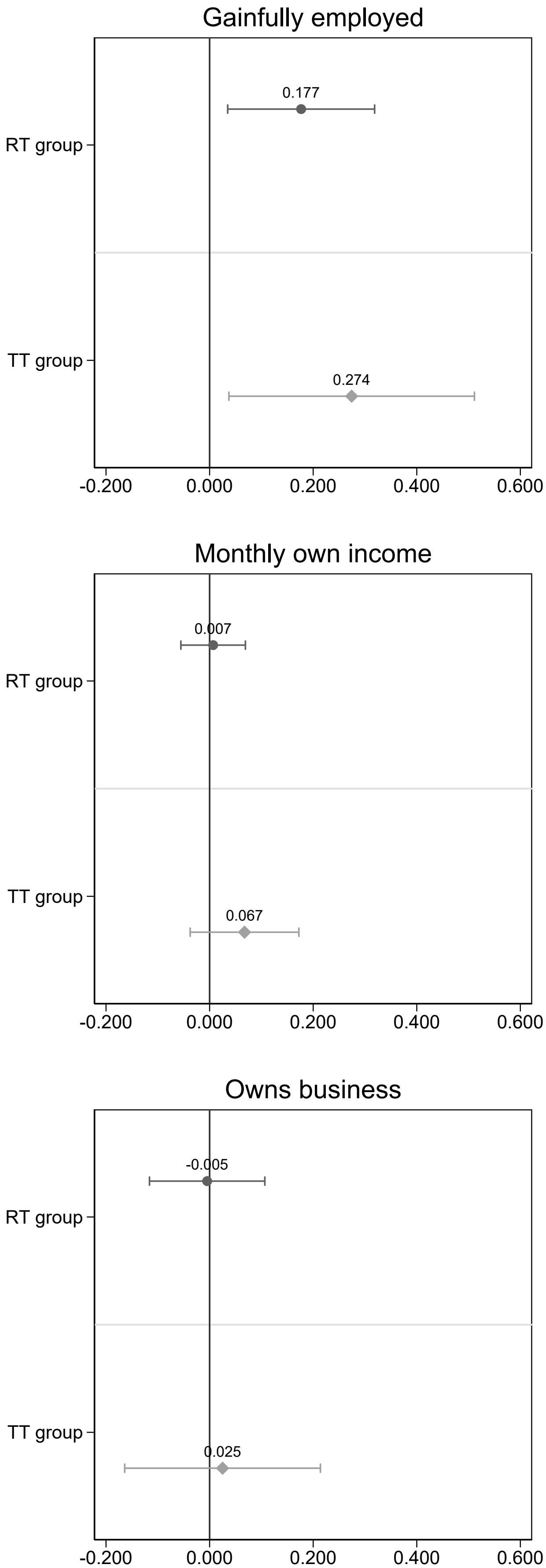

Figure 1: **Group wise standardized effect size (with 95% CI against a one-sided alternative) on selected outcomes**

TABLE 6: **The Romano–Wolf (R–W) multiple hypothesis corrected *p–values* for treatment**

Variable \ **Group**	Random treatment			Trainer treatment		
	Model (1)	Resample (2)	R-W (3)	Model (4)	Resample (5)	R-W (6)
Gainfully employed	0.00	0.00	0.00	0.00	0.00	0.00
Monthly hours worked	0.12	0.14	0.40	0.01	0.01	0.02
Monthly own income	0.88	0.89	1.00	0.37	0.14	0.49
Income working for oneself	0.73	0.67	0.99	0.90	0.84	0.94
Owns business	0.91	0.91	1.00	0.80	0.80	0.94
International Migration	0.94	0.95	1.00	0.21	0.17	0.38
Treatment p-value						
For joint tests		0.00			0.00	

Note: The reported *p-values* refer to H_0: β_1=0 against H_1: β_1>0. The models use district fixed effects. The *p-values* in columns 1 and 4 are generated from simple (uncorrelated) model; the *p-values* in columns 2 and 5 are derived from models that randomly resamples respecting strata and clusters and; the *p-values* in columns 3 and 6 are derived from the Romano-Wolf (R-W) multiple hypotheses corrected models. Romano-Wolf (R-W) *p-values* have been generated using *rwolf* command in Stata, discussed in Clarke et al. (2020). The *p-values* for joint test of significance has been generated using Stata command *randcmd* that conducts a hypothesis test that the treatment (or the group dummy) has no effect, and then tests this hypothesis across equations, relying on bootstrap or randomization inference to calculate the joint distribution of *p-values*; see Young (2019) for more details about the methodology. We have been benefited from the description of the *randcmd* command by David McKenzie that can be found at https://blogs.worldbank.org/en/impactevaluations/overview-multiple-hypothesis-testing-commands-stata.

Table 7: **ITT effect of training by subgroups**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Monthly income working for oneself (4)	Owns business (5)	International Migration (6)
a. Random treatment group						
Male	0.13*	2.92	0.78	0.81	-0.02	-0.02
	(0.07)	(14.42)	(2.42)	(1.73)	(0.05)	(0.06)
Female	0.21	22.95	1.89	-0.43	0.02	0.01
	(0.17)	(19.35)	(1.49)	(0.88)	(0.10)	(0.02)
No education	0.16	28.72	0.36	-2.89	-0.00	0.06*
	(0.15)	(21.19)	(3.84)	(3.76)	(0.05)	(0.03)
Primary education	0.12	4.81	0.91	0.51	-0.10	-0.02
	(0.11)	(14.95)	(3.72)	(2.67)	(0.06)	(0.07)
Secondary education	0.24	-0.68	-0.67	0.70	0.05	-0.04
	(0.20)	(21.76)	(3.88)	(1.93)	(0.09)	(0.18)
Low income	0.24*	19.93	1.51	0.90	0.02	-0.01
	(0.11)	(14.76)	(1.82)	(1.31)	(0.08)	(0.05)
High income	0.07	7.21	-0.37	-0.70	-0.01	0.02
	(0.08)	(15.26)	(3.46)	(2.93)	(0.07)	(0.08)
b. Trainer treatment group						
Male	0.20	27.98	3.63	0.44	-0.10	0.06
	(0.12)	(18.20)	(3.51)	(1.86)	(0.12)	(0.07)
Female	0.29	58.42	1.99	0.43	0.08*	0.08
	(0.35)	(43.72)	(1.88)	(1.14)	(0.04)	(0.05)
No education	0.30	36.62	3.51*	0.80	0.01	0.12
	(0.31)	(39.20)	(1.69)	(0.98)	(0.07)	(0.10)
Primary education	0.14	35.64	0.65	-2.20**	-0.09	0.06
	(0.15)	(26.80)	(1.99)	(0.90)	(0.13)	(0.07)
Secondary education	0.38*	20.02	8.02	2.71	0.20	0.18
	(0.19)	(18.20)	(7.82)	(3.51)	(0.12)	(0.13)
Low income	0.37*	48.06	2.20	0.04	0.08	0.05
	(0.19)	(29.39)	(1.67)	(0.81)	(0.09)	(0.06)
High income	0.11	14.29	1.60	-0.34	-0.14	0.07
	(0.13)	(19.01)	(3.29)	(3.27)	(0.10)	(0.07)

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34, 35-39, and 40-49 years), gender, education (below primary, primary and secondary), marital status, caste, prior training experience and the value of the outcome variable at the baseline as well as district fixed effects while excludes the variable(s) that define the subgroup. SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Study participants reporting nil income in 2014 were considered as low-income.

Appendix A: Additional tables and figures

Table A.1: **Applicants in the follow-up survey by tracking method and trainee type**

	Control (1)	Treatment (2)	All (3)
a. Random treatment group			
Contact in person	100(71.9%)	298(79.9%)	398(77.7%)
Contact over phone	12(8.6%)	32(8.6%)	44(8.6%)
No contact	27(19.4%)	43(11.5%)	70(13.7%)
Total	139(100%)	373(100%)	512(100%)
b. Trainer treatment group			
Contact in person	100(71.9%)	295(78.9%)	395(77.0%)
Contact over phone	12(8.6%)	22(5.9%)	34(6.6%)
No contact	27(19.4%)	57(15.2%)	84(16.4%)
Total	139(100%)	374(100%)	513(100%)
c. Trainer control group			
Contact in person	100(71.9%)	102(68.0%)	202(69.9%)
Contact over phone	12(8.6%)	14(9.3%)	26(9.0%)
No contact	27(19.4%)	34(22.7%)	61(21.1%)
Total	139(100%)	150(100%)	289(100%)
d. All program participants			
Contact in person	100(71.9%)	695(77.5%)	795(76.7%)
Contact over phone	12(8.6%)	68(7.6%)	80(7.7%)
No contact	27(19.4%)	134(14.9%)	161(15.5%)
Grand Total	139(100%)	897(100%)	1,036(100%)

Table A.2: **Difference in the attrited sample by group type: OLS estimate with the main specification**

Group / **Variable**	Random treatment (1)	Trainer treatment (2)	Trainer control (3)
Treatment	-0.08*	-0.13*	-0.08
	(0.04)	(0.07)	(0.08)
Age 20-24	0.03	0.05	0.07
	(0.05)	(0.06)	(0.06)
Age 25-29	-0.05	0.01	-0.01
	(0.06)	(0.08)	(0.08)
Age 30-34	-0.12*	0.01	-0.09
	(0.06)	(0.08)	(0.08)
Age 35-39	-0.03	0.03	0.09
	(0.07)	(0.09)	(0.14)
Age 40-49	-0.10*	-0.09	-0.19
	(0.06)	(0.07)	(0.17)
Female	-0.13***	-0.06	-0.13*
	(0.03)	(0.05)	(0.07)
Education: primary to below SLC	0.00	0.01	0.04
	(0.05)	(0.04)	(0.08)
Education: SLC and beyond	-0.04	-0.05	-0.05
	(0.07)	(0.07)	(0.09)
Never married	-0.01	0.05	-0.00
	(0.05)	(0.05)	(0.06)
Brahmin and Chhetri	0.02	-0.09	0.03
	(0.04)	(0.06)	(0.06)
Prior training participation	-0.08	-0.12***	-0.21***
	(0.06)	(0.04)	(0.06)
Constant	0.28***	0.28**	0.29***
	(0.07)	(0.12)	(0.10)
Adjusted R^2	0.03	0.06	0.07
N	512	513	289

Note: The reference groups are participants aged 15-19 and those having below primary education. SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.3: **Balance during baseline in the attrited sample**

Variable \ **Group**	Random control	Random treatment		Trainer treatment		Trainer control	
	Mean (1)	Mean (2)	Difference (3)	Mean (4)	Difference (5)	Mean (6)	Difference (7)
a. Control variables							
Age 15-19	0.22	0.28	0.06	0.30	0.08	0.24	0.01
	(0.08)	(0.07)	(0.11)	(0.06)	(0.11)	(0.07)	(0.11)
Age 20-24	0.48	0.47	-0.02	0.35	-0.13	0.47	-0.01
	(0.10)	(0.08)	(0.12)	(0.06)	(0.11)	(0.09)	(0.13)
Age 25-29	0.15	0.14	-0.01	0.16	0.01	0.15	-0.00
	(0.07)	(0.05)	(0.09)	(0.05)	(0.09)	(0.06)	(0.09)
Age 30-34	0.00	0.07	0.07	0.16	0.16**	0.12	0.12*
	(0.00)	(0.04)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)
Age 35-39	0.15	0.05	-0.10	0.04	-0.11*	0.03	-0.12*
	(0.07)	(0.03)	(0.07)	(0.02)	(0.06)	(0.03)	(0.07)
Age 40-49	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.11	0.16	0.05	0.25	0.13	0.09	-0.02
	(0.06)	(0.06)	(0.09)	(0.06)	(0.09)	(0.05)	(0.08)
Education: below primary	0.19	0.16	-0.02	0.14	-0.04	0.06	-0.13
	(0.08)	(0.06)	(0.09)	(0.05)	(0.09)	(0.04)	(0.08)
Education: primary to below SLC	0.56	0.49	-0.07	0.60	0.04	0.62	0.06
	(0.10)	(0.08)	(0.12)	(0.07)	(0.12)	(0.08)	(0.13)
Education: SLC and beyond	0.26	0.35	0.09	0.26	0.00	0.32	0.06
	(0.09)	(0.07)	(0.12)	(0.06)	(0.10)	(0.08)	(0.12)
Never married	0.48	0.51	0.03	0.56	0.08	0.59	0.11
	(0.10)	(0.08)	(0.12)	(0.07)	(0.12)	(0.09)	(0.13)
Brahmin and Chhetri	0.22	0.26	0.03	0.12	-0.10	0.26	0.04
	(0.08)	(0.07)	(0.11)	(0.04)	(0.08)	(0.08)	(0.11)
Prior training participation	0.00	0.05	0.05	0.07	0.07	0.03	0.03
	(0.00)	(0.03)	(0.04)	(0.03)	(0.05)	(0.03)	(0.03)
F-test (p-value)	-	-	0.74	-	0.25	-	0.69
Observations	[27]	[43]	[70]	[57]	[84]	[34]	[61]
b. Outcome variables							
Gainfully employed	0.30	0.35	0.05	0.23	-0.07	0.21	-0.09
	(0.09)	(0.07)	(0.12)	(0.06)	(0.10)	(0.07)	(0.11)
Monthly hours worked	108.67	100.35	-8.32	83.28	-25.39	50.15	-58.52**
	(19.72)	(17.67)	(27.24)	(13.21)	(23.50)	(12.90)	(22.75)
Monthly own income	1.92	2.23	0.31	1.70	-0.23	1.13	-0.79
	(0.57)	(0.53)	(0.81)	(0.56)	(0.90)	(0.39)	(0.68)
Income working for oneself	0.55	0.98	0.43	0.54	-0.01	0.38	-0.17
	(0.31)	(0.36)	(0.52)	(0.44)	(0.68)	(0.27)	(0.42)
Owns business	0.07	0.12	0.04	0.04	-0.04	0.03	-0.04
	(0.05)	(0.05)	(0.07)	(0.02)	(0.05)	(0.03)	(0.06)
International Migration	0.07	0.07	-0.00	0.02	-0.06	0.15	0.07
	(0.05)	(0.04)	(0.06)	(0.02)	(0.04)	(0.06)	(0.08)
F-test (p-value)	-	-	0.98	-	0.77	-	0.27
Observations	[27]	[43]	[70]	[57]	[84]	[34]	[61]

Note: Means are reported; SEs are in the parentheses. Column 3 shows the difference between RT and the RC group; column 5 shows the same between TT and RC group and column 7 shows the same between TC and RC group. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The *F-test* of joint significance runs a regression of treatment on all the control/outcome variables in the groups and then tests the null hypothesis that all the slope coefficients are zero.

Table A.4: **After training mean outcomes by group type and their differences with RC group**

Variable \ Group	Random control	Random treatment		Trainer treatment		Trainer control	
	Mean	Mean	Difference	Mean	Difference	Mean	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gainfully employed	0.61	0.78	0.17***	0.80	0.19***	0.54	-0.06
	(0.05)	(0.02)	(0.05)	(0.02)	(0.05)	(0.04)	(0.06)
Monthly hours worked	171.07	185.43	14.36	188.32	17.25*	153.06	-18.01
	(8.63)	(4.69)	(9.58)	(4.97)	(9.89)	(9.14)	(12.58)
Monthly own income	8.46	9.48	1.02	9.75	1.29	4.87	-3.59**
	(1.41)	(1.01)	(1.93)	(1.11)	(2.05)	(0.78)	(1.60)
Income working for oneself	3.07	3.23	0.16	3.38	0.31	1.50	-1.57
	(1.21)	(0.96)	(1.80)	(0.71)	(1.39)	(0.52)	(1.31)
Owns business	0.14	0.15	0.01	0.24	0.09**	0.12	-0.02
	(0.03)	(0.02)	(0.04)	(0.02)	(0.04)	(0.03)	(0.04)
International Migration	0.15	0.17	0.02	0.15	0.00	0.20	0.05
	(0.03)	(0.02)	(0.04)	(0.02)	(0.04)	(0.04)	(0.05)
F-test (p-value)	-	-	0.10	-	0.00	-	0.22
Observations	[139]	[373]	[419]	[374]	[404]	[150]	[208]

Note: Means are reported; SEs are in the parentheses. Column 3 shows the difference between RT and the RC group; column 5 shows the same between TT and RC group and column 7 shows the same between TC and RC group. The *s indicate the significance in difference in means using *t-tests*: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The *F-test* of joint significance runs a regression of treatment on all the outcome variables in the groups and then tests the null hypothesis that all the slope coefficients are zero.

Table A.5: **ITT effect on random treatment group**
(On some other outcome variables)

	Gainfully employed (1)	Average hours worked (2)	Internal migration (3)	Has a formal business (4)	Other family members' income (5)
a. With district fixed effects					
Treatment	0.09	0.48	0.07**	-0.01	1.07
	(0.07)	(0.36)	(0.03)	(0.01)	(1.40)
Age 20-24	0.10	-0.43	0.03	-0.00	-0.55
	(0.07)	(0.73)	(0.07)	(0.01)	(1.07)
Age 25-29	0.14	-0.40	0.02	0.01	-1.03
	(0.08)	(0.88)	(0.08)	(0.01)	(2.59)
Age 30-34	0.14*	-0.06	0.04	-0.00	1.49
	(0.07)	(0.97)	(0.08)	(0.00)	(2.78)
Age 35-39	0.14	-0.78	0.03	0.03	2.87
	(0.08)	(0.95)	(0.09)	(0.02)	(3.58)
Age 40-49	0.27*	-0.86	0.01	-0.01	-3.73
	(0.15)	(1.29)	(0.07)	(0.01)	(2.39)
Female	-0.29***	-0.50	-0.08	0.02	5.49
	(0.07)	(0.67)	(0.06)	(0.02)	(3.16)
Education: primary	0.09**	-0.36	0.02	0.01	4.73*
to below SLC	(0.03)	(0.25)	(0.02)	(0.01)	(2.55)
Education: SLC and	0.02	0.23	0.00	0.02	10.15**
beyond	(0.03)	(0.40)	(0.06)	(0.02)	(3.82)
Never married	0.04	-0.91**	0.03	-0.02	-2.78**
	(0.06)	(0.43)	(0.05)	(0.01)	(1.16)
Brahmin and Chhetri	-0.01	-0.37	0.00	-0.01	2.12
	(0.03)	(0.31)	(0.05)	(0.01)	(3.09)
Prior training	-0.05	-0.53	0.06	-0.01	-3.41
participation	(0.10)	(0.53)	(0.07)	(0.00)	(2.02)
Y_{t-1}	0.12**	0.09	-0.03	0.99***	0.15**
	(0.05)	(0.07)	(0.07)	(0.01)	(0.06)
Constant	0.65***	6.36***	0.02	-0.00	-1.19
	(0.10)	(0.76)	(0.09)	(0.01)	(3.58)
Adjusted R^2	0.21	0.12	0.10	0.33	0.12
N	461	419	398	442	442
b. With regression adjustment					
Treatment	0.10***	0.46	0.08***	-0.01	1.19
	(0.04)	(0.32)	(0.02)	(0.01)	(1.26)
N	461	419	398	442	442
c. With inverse probability weighting					
Treatment	0.10***	0.46	0.08***	-0.01	1.19
	(0.04)	(0.32)	(0.02)	(0.01)	(1.26)
N	461	419	398	442	442
d. With selection of covariates using Lasso					
Treatment	0.10**	0.48	0.09***	-0.00	1.28
	(0.04)	(0.33)	(0.02)	(0.01)	(1.27)
N	461	419	398	442	442
e. Lee bounds					
lower	0.10**	-0.19	-0.02	-0.01	-2.41
	(0.05)	(0.39)	(0.06)	(0.01)	(1.49)
upper	0.23***	1.05***	0.09***	-0.00	2.22
	(0.06)	(0.40)	(0.03)	(0.01)	(1.50)
N	512	512	512	512	512
f. Significance level with randomization inference					
Treatment	0.09***	0.48	0.07**	-0.01	1.07
N	461	419	398	442	442

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. The definition of the variable "Gainfully employed" includes home cultivation, a proxy for subsistence farming.

Table A.6: **ITT effect on the trainer treatment group**
(On some other outcome variables)

	Gainfully employed (1)	Average hours worked (2)	Internal migration (3)	Has a formal business (4)	Other family members' income (5)
a. With district fixed effects					
Treatment	0.18	1.23*	0.05	0.00	-2.19
	(0.12)	(0.63)	(0.03)	(0.01)	(1.59)
Age 20-24	0.07	-0.33	-0.07	-0.00	-1.70
	(0.05)	(0.46)	(0.05)	(0.01)	(1.38)
Age 25-29	0.08*	-0.35	-0.08	0.01	-0.72
	(0.04)	(0.51)	(0.07)	(0.01)	(1.67)
Age 30-34	0.12	0.12	-0.12	-0.02	-3.08
	(0.07)	(0.60)	(0.09)	(0.02)	(2.37)
Age 35-39	0.19**	0.47	-0.10	-0.02	-3.19
	(0.07)	(0.67)	(0.07)	(0.03)	(1.98)
Age 40-49	0.07	3.05**	-0.09	0.16	27.60
	(0.16)	(1.26)	(0.06)	(0.15)	(26.11)
Female	-0.25***	-0.00	-0.05**	-0.00	1.22
	(0.07)	(0.36)	(0.02)	(0.02)	(2.75)
Education: primary to below SLC	0.01	0.02	-0.08	0.00	1.35
	(0.05)	(0.44)	(0.06)	(0.01)	(1.19)
Education: SLC and beyond	-0.01	-0.18	-0.08	0.01	5.25**
	(0.06)	(0.63)	(0.06)	(0.01)	(2.03)
Never married	-0.01	-0.33	0.01	-0.03	-2.81
	(0.04)	(0.34)	(0.06)	(0.03)	(1.83)
Brahmin and Chhetri	-0.06	0.68	0.03	0.01	-0.42
	(0.05)	(0.59)	(0.06)	(0.01)	(1.88)
Prior training participation	0.07	-0.31	0.02	0.00	-1.19
	(0.06)	(0.67)	(0.05)	(0.01)	(1.16)
Y_{t-1}	0.03	0.02	0.22*		0.07
	(0.04)	(0.05)	(0.11)		(0.06)
Constant	0.73***	5.24***	0.17**	0.02	7.63**
	(0.12)	(1.01)	(0.08)	(0.02)	(3.64)
Adjusted R^2	0.17	0.11	0.14	0.14	0.16
N	453	404	395	429	429
b. With regression adjustment					
Treatment	0.11***	0.58*	0.08***	-0.00	1.15
	(0.04)	(0.33)	(0.02)	(0.01)	(1.08)
N	453	404	395	429	429
c. With inverse probability weighting					
Treatment	0.11***	0.58*	0.08***	-0.00	1.15
	(0.04)	(0.33)	(0.02)	(0.01)	(1.08)
N	453	404	395	429	429
d. With selection of covariates using Lasso					
Treatment	0.10***	0.58*	0.08***	0.00	0.88
	(0.04)	(0.33)	(0.02)	(0.01)	(1.20)
N	453	404	395	429	429
e. Lee bounds					
lower	0.09*	0.10	-0.01	-0.01	-1.29
	(0.05)	(0.41)	(0.06)	(0.01)	(1.53)
upper	0.19***	0.98**	0.09***	0.00	1.32
	(0.06)	(0.43)	(0.03)	(0.01)	(1.43)
N	513	513	513	513	513
f. Significance level with randomization inference					
Treatment	0.18***	1.23***	0.05	0.00	-2.19
N	453	404	395	429	429

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The definition of the variable "Gainfully employed" includes home cultivation, a proxy for subsistence farming.

Table A.7: **Changes of the trainer control group**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. With district fixed effects						
Treatment	-0.04	-6.78	-1.48	-0.57	-0.04	0.09
	(0.08)	(19.78)	(1.53)	(0.72)	(0.04)	(0.06)
b. With regression adjustment						
Treatment	-0.09	-20.58	-3.40**	-1.21	-0.01	0.02
	(0.06)	(12.58)	(1.66)	(1.36)	(0.04)	(0.05)
c. With inverse probability weighting						
Treatment	-0.09	-20.58	-3.40**	-1.21	-0.01	0.02
	(0.06)	(12.58)	(1.66)	(1.36)	(0.04)	(0.05)
d. With selection of covariates using Lasso						
Treatment	-0.09	-18.01	-3.59**	-1.57	-0.02	0.05
	(0.06)	(12.51)	(1.60)	(1.32)	(0.04)	(0.05)
e. Lee bounds						
lower	-0.11	-23.87	-4.56**	-1.76	-0.02	-0.00
	(0.07)	(17.84)	(1.97)	(1.49)	(0.05)	(0.06)
upper	-0.09	-14.81	-2.93	-0.21	-0.00	0.03
	(0.08)	(19.05)	(2.63)	(1.96)	(0.07)	(0.05)
f. With randomization inference						
Treatment	-0.04	-6.78	-1.48	-0.57	-0.04	0.09*
g. Random control mean						
At endline	0.61	171.07	8.46	3.07	0.14	0.15
N	239	208	228	228	228	239

Note: See the notes in Table 3.

Table A.8: **ITT effect on the trainer control group**
(On some other outcome variables)

	Gainfully employed (1)	Average hours worked (2)	Internal migration (3)	Has a formal business (4)	Other family members' income (5)
a. With district fixed effects					
Treatment	0.02	-0.23	0.01	0.02	-3.39
	(0.07)	(0.66)	(0.03)	(0.02)	(2.18)
Age 20-24	0.13	-0.36	0.01	0.02	3.45
	(0.09)	(0.98)	(0.07)	(0.02)	(2.76)
Age 25-29	0.22**	-0.28	-0.01	0.04	1.82
	(0.09)	(0.79)	(0.05)	(0.04)	(3.02)
Age 30-34	0.12	-0.59	-0.03	-0.00	0.98
	(0.12)	(1.13)	(0.07)	(0.03)	(2.78)
Age 35-39	0.24**	0.11	-0.02	0.01	1.81
	(0.11)	(1.29)	(0.06)	(0.03)	(2.63)
Age 40-49	0.30**	1.84	-0.00	0.00	6.00
	(0.13)	(1.44)	(0.07)	(0.02)	(3.51)
Female	-0.27***	-2.02**	-0.06*	0.00	-0.09
	(0.05)	(0.72)	(0.03)	(0.04)	(1.91)
Education: primary to below SLC	0.08	-0.39	0.01	0.01	-1.57
	(0.05)	(0.56)	(0.02)	(0.04)	(2.01)
Education: SLC and beyond	-0.03	-0.90	0.04	0.01	0.26
	(0.06)	(0.86)	(0.05)	(0.01)	(2.05)
Never married	0.08	-0.05	0.00	-0.01	-0.17
	(0.09)	(0.55)	(0.03)	(0.03)	(2.69)
Brahmin and Chhetri	0.02	-0.02	-0.03	-0.03*	-0.21
	(0.09)	(0.80)	(0.02)	(0.02)	(1.99)
Prior training participation	0.04	-0.18	0.12	-0.02	-1.04
	(0.08)	(0.75)	(0.08)	(0.02)	(2.27)
Y_{t-1}	0.15**	0.13	0.12		0.21*
	(0.06)	(0.10)	(0.21)		(0.11)
Constant	0.55***	6.41***	0.04	-0.00	4.49*
	(0.12)	(1.06)	(0.06)	(0.03)	(2.42)
Adjusted R^2	0.24	0.10	-0.03	-0.05	0.12
N	239	208	202	228	228
b. With regression adjustment					
Treatment	-0.06	-0.69	0.01	0.02	-0.49
	(0.05)	(0.42)	(0.03)	(0.02)	(1.30)
N	239	208	202	228	228
c. With inverse probability weighting					
Treatment	-0.06	-0.69	0.01	0.02	-0.49
	(0.05)	(0.42)	(0.03)	(0.02)	(1.30)
N	239	208	202	228	228
d. With selection of covariates using Lasso					
Treatment	-0.05	-0.60	0.02	0.02	-0.97
	(0.05)	(0.42)	(0.03)	(0.02)	(1.28)
N	239	208	202	228	228
e. Lee bounds					
lower	-0.02	-0.89	0.02	0.02	-1.24
	(0.06)	(0.56)	(0.03)	(0.02)	(1.45)
upper	-0.01	-0.26	0.05**	0.03*	0.87
	(0.07)	(0.59)	(0.02)	(0.01)	(1.80)
N	289	289	289	289	289
f. Significance level with randomization inference					
Treatment	0.02	-0.23	0.01	0.02	-3.39**
N	239	208	202	228	228

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The definition of the variable "Gainfully employed" includes home cultivation, a proxy for subsistence farming.

Table A.9: **Pure treatment effect on the trainer treatment group**
(Corresponds to Table 4)
(Compares TT group with the part of RC group members whose score is above the minimum score of the comparable (category-wise) TT group)

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. With district fixed effects						
Treatment	0.26	39.56**	1.23	-1.21	-0.03	0.07*
	(0.16)	(17.85)	(1.50)	(0.98)	(0.09)	(0.04)
b. With regression adjustment						
Treatment	0.24***	29.16***	3.71***	1.49	0.09*	0.05
	(0.06)	(10.61)	(1.41)	(0.95)	(0.04)	(0.04)
c. With inverse probability weighting						
Treatment	0.24***	29.16***	3.71***	1.49	0.09*	0.05
	(0.06)	(10.61)	(1.41)	(0.95)	(0.04)	(0.04)
d. With selection of covariates using Lasso						
Treatment	0.23***	27.41**	3.73**	0.98	0.09**	0.06
	(0.06)	(11.22)	(1.47)	(0.86)	(0.04)	(0.04)
e. Lee bounds						
lower	0.21***	27.00*	3.56**	1.79*	0.11**	-0.01
	(0.06)	(14.12)	(1.63)	(0.96)	(0.05)	(0.06)
upper	0.28***	28.56*	3.36	1.73	0.11*	0.06
	(0.08)	(16.16)	(2.24)	(1.59)	(0.06)	(0.04)
f. Significance level with randomization inference						
Treatment	0.26***	39.56***	1.23	-1.21	-0.03	0.07*
g. Random control mean						
At endline	0.57	161.87	6.11	1.62	0.12	0.10
N	407	366	385	385	385	407

Note: See the notes in Table 3.

TABLE A.10: **Pure selection effect**
(Corresponds to Table A.8)
(Divides the RC group into two parts-one with members having category-wise above-mean scores while the reference group includes those with category-wise below-mean scores)

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. With district fixed effects						
Treatment	-0.02	-33.37*	-7.87*	-3.04	-0.03	-0.07
	(0.14)	(16.44)	(4.21)	(2.73)	(0.12)	(0.10)
b. With regression adjustment						
Treatment	-0.01	-18.22	-4.37*	-2.30	-0.07	-0.03
	(0.09)	(16.95)	(2.34)	(2.02)	(0.06)	(0.06)
c. With inverse probability weighting						
Treatment	-0.01	-18.22	-4.37*	-2.30	-0.07	-0.03
	(0.09)	(16.95)	(2.34)	(2.02)	(0.06)	(0.06)
d. With selection of covariates using Lasso						
Treatment	-0.07	-26.54	-6.03**	-1.75	-0.04	-0.08
	(0.09)	(16.98)	(2.68)	(2.34)	(0.06)	(0.07)
e. Lee bounds						
lower	-0.12	-46.44**	-6.30*	-2.86	-0.06	-0.07
	(0.14)	(20.02)	(3.66)	(2.48)	(0.08)	(0.08)
upper	0.09	8.09	2.33	1.55	0.11	0.09
	(0.12)	(18.91)	(2.00)	(1.40)	(0.07)	(0.09)
f. Significance level with randomization inference						
Treatment	-0.02	-33.37*	-7.87***	-3.04	-0.03	-0.07
g. Random control mean						
At endline	0.64	184.21	11.37	4.19	0.16	0.19
N	114	103	112	112	112	114

Note: See the notes in Table 3.

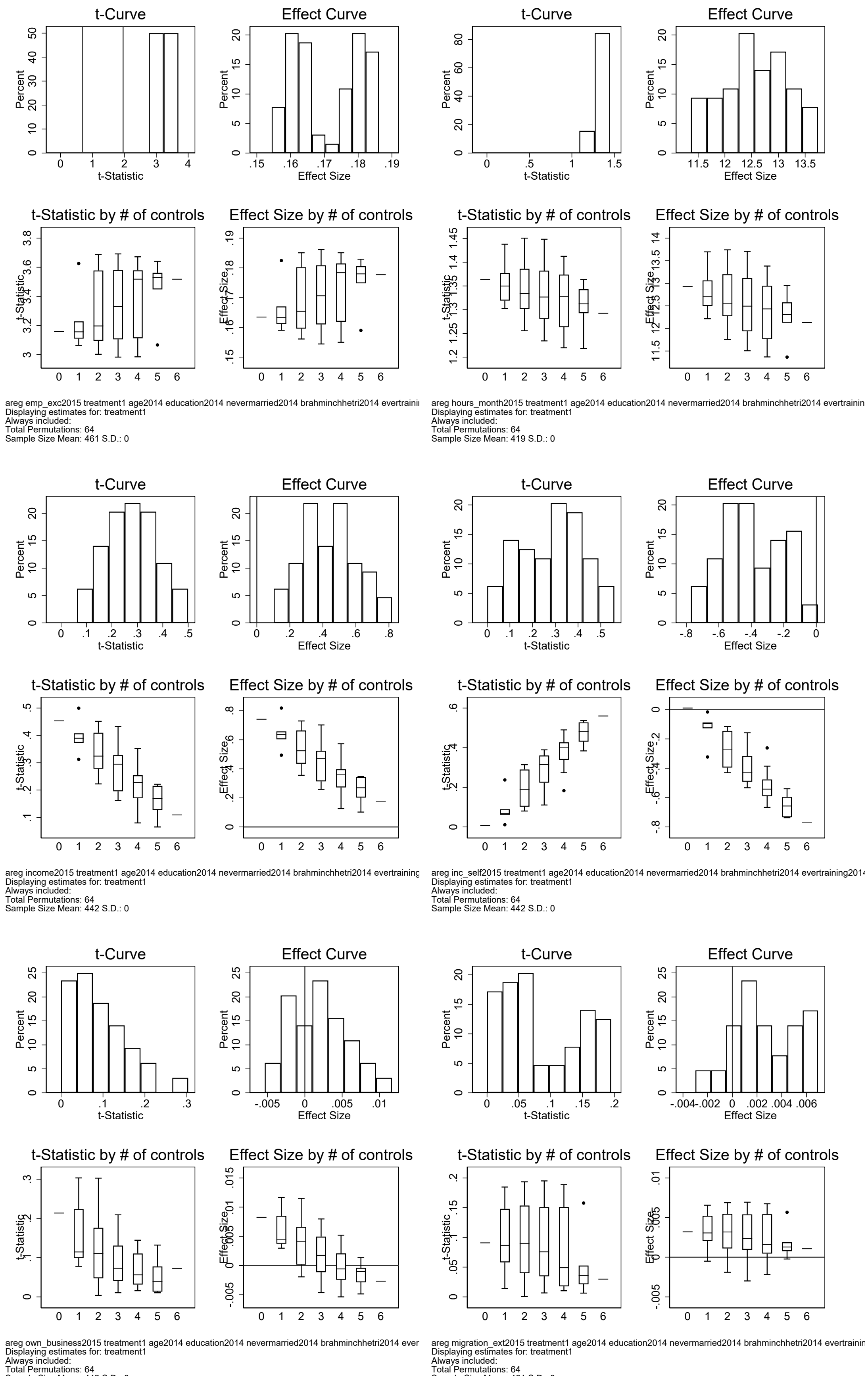

FIGURE A.1: **Specification check for p-hacking for the random treatment group**

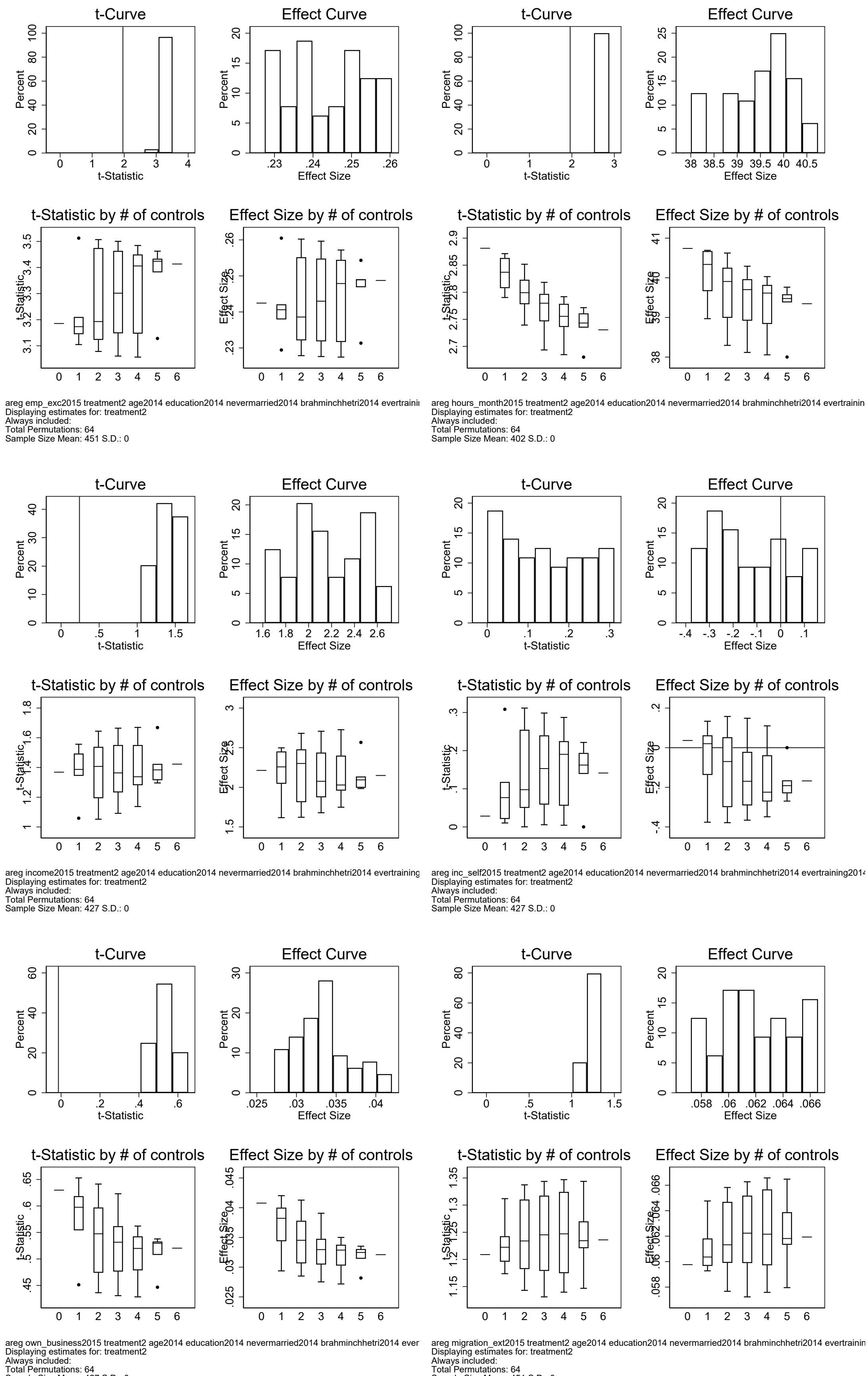

FIGURE A.2: **Specification check for p-hacking for the trainer treatment group**

Table A.11: **Minimum detectable effect size (MDES) by outcomes and group type**

	Actual mean			Sample size	
	Control (1)	Treatment (2)	MDES (3)	Control (4)	Treatment (5)
a. Random treatment group					
Gainfully employed	0.61	0.78	0.14	114	347
Monthly hours worked	171.07	185.43	27.69	103	316
Monthly own income	8.46	9.48	4.89	112	330
Income working for oneself	3.07	3.23	4.36	112	330
Owns business	0.14	0.15	0.11	112	330
International Migration	0.15	0.17	0.11	114	347
b. Trainer treatment group					
Gainfully employed	0.61	0.80	0.14	114	339
Monthly hours worked	171.07	188.32	28.07	103	301
Monthly own income	8.46	9.75	5.05	112	317
Income working for oneself	3.07	3.38	3.95	112	317
Owns business	0.14	0.24	0.12	112	317
International Migration	0.15	0.15	0.11	114	339

Note: The MDESs are based on observed standard deviations in the actual outcomes for the control group and the actual sample size of the treatment and control groups. For binary outcomes, the MDES is expressed in terms of proportions. We assume 80% power and a two-sided test at a significance level of 5 percent.

Table A.12: **LATE of training by groups**
(Corresponds to Table 5)

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. With separate dummies for the two treatment groups						
Trained–RT group	0.20**	15.69	-0.35	-1.04	-0.01	0.01
	(0.09)	(11.12)	(1.37)	(0.95)	(0.06)	(0.05)
Trained–TT group	0.31***	33.26**	4.20	1.83	0.06	0.06
	(0.10)	(14.03)	(3.32)	(2.31)	(0.10)	(0.07)
b. With interaction dummies for the two Treatment groups						
Trained	0.22**	17.67	-0.40	-1.16	-0.01	0.01
	(0.09)	(12.66)	(1.53)	(1.07)	(0.07)	(0.05)
Trained × TT	0.12**	17.93**	4.54	2.84	0.07	0.05
	(0.05)	(8.78)	(3.34)	(2.74)	(0.08)	(0.08)
N	800	720	759	759	759	800

Note: The models additionally include district fixed effects. SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B: Tables related to the tests for heterogeneity

Table B.1: **ITT effect on RT group by sex**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Treatment (male)	0.13*	2.92	0.78	0.81	-0.02	-0.02
	(0.07)	(14.42)	(2.42)	(1.73)	(0.05)	(0.06)
Age 20-24	-0.02	-16.37	1.17	-0.50	-0.01	0.01
	(0.06)	(27.03)	(2.09)	(0.87)	(0.04)	(0.05)
Age 25-29	-0.10	-34.01	1.12	3.02	0.04	-0.10
	(0.11)	(33.54)	(3.93)	(3.77)	(0.07)	(0.11)
Age 30-34	-0.21*	-8.92	3.74	6.99	0.12	-0.09
	(0.10)	(35.61)	(4.73)	(4.84)	(0.08)	(0.13)
Age 35-39	-0.36**	-63.64*	0.34	5.61*	0.17	-0.15
	(0.16)	(33.97)	(3.90)	(3.18)	(0.14)	(0.14)
Age 40-49	-0.02	-18.71	-1.23	-2.39	-0.05	-0.07
	(0.12)	(33.10)	(3.33)	(3.23)	(0.08)	(0.13)
Education: primary	-0.03	-37.97**	1.92	3.42	0.06	-0.07
to below SLC	(0.06)	(12.79)	(2.38)	(2.02)	(0.06)	(0.10)
Education: SLC and	0.05	-3.55	2.03	3.90	0.15	0.01
beyond	(0.07)	(15.35)	(3.02)	(2.54)	(0.10)	(0.10)
Never married	-0.16*	-23.05	-1.06	1.41	-0.04	0.07
	(0.08)	(16.01)	(3.30)	(2.87)	(0.06)	(0.12)
Brahmin and Chhetri	-0.08	-5.92	2.58	3.79	0.03	-0.01
	(0.05)	(10.75)	(4.40)	(4.31)	(0.04)	(0.10)
Prior training	-0.16	-31.74*	-2.60	-0.75	0.04	-0.10
participation	(0.16)	(16.71)	(2.57)	(2.21)	(0.09)	(0.07)
Y_{t-1}	0.11	0.09	0.19*	-0.06	0.11	0.15
	(0.07)	(0.09)	(0.10)	(0.06)	(0.09)	(0.19)
Constant	0.85***	225.13***	7.18	-4.14	0.01	0.32**
	(0.09)	(25.20)	(6.04)	(5.76)	(0.11)	(0.13)
Adjusted R^2	0.12	0.17	0.22	0.33	0.13	0.18
N	291	253	272	272	272	291
Treatment (female)	0.21	22.95	1.89	-0.43	0.02	0.01
	(0.17)	(19.35)	(1.49)	(0.88)	(0.10)	(0.02)
Age 20-24	0.07	-12.24	-2.89	-5.74	0.04	-0.04
	(0.18)	(31.92)	(3.56)	(3.88)	(0.06)	(0.04)
Age 25-29	0.27	5.89	-1.35	-3.83	0.18*	-0.04
	(0.25)	(40.14)	(4.29)	(3.76)	(0.09)	(0.04)
Age 30-34	0.15	-5.77	5.40	2.96	0.15	-0.03
	(0.24)	(38.29)	(5.38)	(4.85)	(0.10)	(0.03)
Age 35-39	0.14	12.10	-3.31	-5.20	0.08	-0.03
	(0.24)	(28.90)	(4.20)	(4.31)	(0.07)	(0.03)
Age 40-49	0.43	-33.21	-0.70	-3.13	0.02	-0.03
	(0.26)	(68.62)	(3.21)	(2.00)	(0.05)	(0.03)
Education: primary	0.18*	8.71	1.11	2.01	0.21**	0.01
to below SLC	(0.09)	(16.80)	(1.65)	(1.71)	(0.08)	(0.02)
Education: SLC and	0.16	7.00	9.97	9.48	0.20	0.00
beyond	(0.14)	(22.03)	(7.02)	(6.78)	(0.13)	(0.01)
Never married	0.09	-32.77	-3.65	-6.77*	-0.20*	0.03
	(0.24)	(18.71)	(2.87)	(3.49)	(0.09)	(0.03)
Brahmin and Chhetri	0.01	12.52	5.38	7.75	0.14*	-0.01
	(0.09)	(15.29)	(3.59)	(4.45)	(0.07)	(0.01)
Prior training	-0.11	12.52	6.33**	6.80**	0.01	0.05
participation	(0.18)	(17.51)	(2.50)	(2.76)	(0.05)	(0.06)
Y_{t-1}	0.25**	0.06	-0.16	-0.20**	0.03	0.02
	(0.09)	(0.09)	(0.11)	(0.09)	(0.16)	(0.02)
Constant	0.18	154.67***	2.35	3.97	0.00	0.02
	(0.25)	(42.89)	(3.22)	(2.25)	(0.10)	(0.02)
Adjusted R^2	0.17	0.15	0.10	0.19	0.34	0.21
N	170	166	170	170	170	170

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: **ITT effect on TT group by sex**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Treatment (male)	0.20	27.98	3.63	0.44	-0.10	0.06
	(0.12)	(18.20)	(3.51)	(1.86)	(0.12)	(0.07)
Age 20-24	0.02	-13.80	3.54	-1.47	-0.08	0.05
	(0.07)	(17.40)	(2.12)	(1.36)	(0.05)	(0.05)
Age 25-29	0.03	3.79	2.96	1.66	-0.02	-0.04
	(0.09)	(24.91)	(4.42)	(2.04)	(0.05)	(0.07)
Age 30-34	0.05	-1.83	0.33	3.27	0.09	-0.18*
	(0.08)	(28.56)	(4.39)	(2.01)	(0.11)	(0.09)
Age 35-39	-0.04	10.63	0.83	3.57**	0.23*	-0.03
	(0.13)	(31.01)	(3.44)	(1.65)	(0.13)	(0.06)
Age 40-49	0.04	181.20***	66.41	73.00	0.24	0.15
	(0.07)	(62.51)	(48.55)	(56.36)	(0.23)	(0.35)
Education: primary to below SLC	-0.12**	-11.20	-0.20	0.39	-0.07	0.11*
	(0.06)	(19.38)	(3.65)	(4.17)	(0.09)	(0.05)
Education: SLC and beyond	-0.07	-13.36	5.25	2.62	0.00	0.07
	(0.09)	(29.54)	(4.95)	(4.16)	(0.10)	(0.08)
Never married	-0.16	-6.74	-5.53	-0.50	-0.01	-0.12*
	(0.10)	(15.84)	(4.31)	(1.60)	(0.05)	(0.07)
Brahmin and Chhetri	-0.06	12.94	3.73	1.13	0.04	-0.10**
	(0.08)	(24.09)	(2.48)	(1.90)	(0.06)	(0.05)
Prior training participation	-0.01	-6.70	-3.30	-0.49	-0.01	-0.06
	(0.07)	(21.01)	(2.17)	(1.21)	(0.05)	(0.07)
Y_{t-1}	0.05	0.07	-0.05	-0.00	0.21**	-0.06
	(0.06)	(0.07)	(0.12)	(0.10)	(0.08)	(0.14)
Constant	0.80***	162.68***	7.90	0.51	0.20	0.21**
	(0.17)	(44.71)	(6.02)	(3.80)	(0.16)	(0.08)
Adjusted R^2	0.09	0.12	0.17	0.40	0.19	0.03
N	267	226	246	246	246	267
Treatment (female)	0.29	58.42	1.99	0.43	0.08*	0.08
	(0.35)	(43.72)	(1.88)	(1.14)	(0.04)	(0.05)
Age 20-24	-0.06	-4.66	0.00	0.10	0.04	0.06
	(0.11)	(34.83)	(1.55)	(1.65)	(0.09)	(0.04)
Age 25-29	-0.06	-21.21	-3.06	-3.39	-0.06	0.06
	(0.13)	(26.11)	(3.22)	(3.33)	(0.09)	(0.04)
Age 30-34	-0.05	-11.30	-1.33	-1.39	0.01	0.06*
	(0.14)	(27.69)	(2.72)	(2.96)	(0.09)	(0.04)
Age 35-39	0.11	13.44	-2.18	-3.85	-0.02	0.02
	(0.14)	(49.04)	(3.80)	(3.29)	(0.09)	(0.04)
Age 40-49	-0.02	12.86	-6.01	-4.89	-0.13	0.02
	(0.28)	(45.41)	(4.69)	(3.90)	(0.10)	(0.03)
Education: primary to below SLC	-0.12	7.56	-1.02	-0.77	0.05	-0.02
	(0.08)	(14.81)	(2.29)	(2.38)	(0.07)	(0.01)
Education: SLC and beyond	-0.16	-15.38	-3.27	-3.66	-0.11	-0.01
	(0.11)	(20.63)	(2.38)	(2.34)	(0.13)	(0.03)
Never married	0.10	-3.26	-1.60	-1.70	-0.01	0.03
	(0.08)	(27.47)	(2.47)	(2.85)	(0.13)	(0.03)
Brahmin and Chhetri	0.14	43.62*	5.22	4.59*	0.18*	0.04
	(0.14)	(23.38)	(3.02)	(2.30)	(0.10)	(0.05)
Prior training participation	0.24*	22.36	6.48	5.43	0.11	-0.07**
	(0.13)	(34.70)	(4.46)	(4.71)	(0.12)	(0.03)
Y_{t-1}	0.12	-0.02	0.09*	0.06	0.11	
	(0.18)	(0.07)	(0.04)	(0.06)	(0.18)	
Constant	0.46	150.69**	4.79*	4.61	0.22**	-0.08
	(0.30)	(61.53)	(2.51)	(2.77)	(0.09)	(0.06)
Adjusted R^2	0.17	0.13	0.12	0.14	0.34	-0.05
N	186	178	183	183	183	186

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table B.3: **ITT effect on RT group by educational groups**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Treatment (no education)	0.16	28.72	0.36	-2.89	-0.00	0.06*
	(0.15)	(21.19)	(3.84)	(3.76)	(0.05)	(0.03)
Age 20-24	0.29	44.59***	6.07	-2.65	0.01	0.23*
	(0.25)	(9.06)	(4.49)	(3.52)	(0.04)	(0.12)
Age 25-29	0.45	42.31**	8.03**	0.00	0.11*	0.22
	(0.29)	(16.74)	(3.64)	(1.95)	(0.06)	(0.13)
Age 30-34	0.31	22.61	9.21**	2.28	0.08*	0.26*
	(0.26)	(17.91)	(3.79)	(2.42)	(0.04)	(0.15)
Age 35-39	0.33	30.50	2.95	-2.68	0.07	0.23
	(0.29)	(27.46)	(3.26)	(2.97)	(0.05)	(0.15)
Age 40-49	0.62*	64.77***	8.01*	-1.48	0.04	0.24*
	(0.34)	(15.20)	(3.80)	(2.42)	(0.03)	(0.13)
Female	-0.13	-52.82*	-21.99*	-13.95	-0.07	-0.17
	(0.21)	(28.49)	(11.72)	(12.29)	(0.08)	(0.11)
Never married	0.52	93.61***	8.34	-2.21	0.05	0.35*
	(0.36)	(28.56)	(5.12)	(3.53)	(0.04)	(0.17)
Brahmin and Chhetri	-0.12	-5.39	-0.43	1.65	-0.03	0.01
	(0.21)	(21.57)	(1.58)	(1.93)	(0.03)	(0.03)
Prior training participation	-0.27*	-22.29	14.46*	15.26*	0.14*	0.04
	(0.14)	(52.36)	(7.51)	(8.55)	(0.07)	(0.05)
Y_{t-1}	0.21**	0.00	0.19	-0.66***	0.06	-0.04
	(0.08)	(0.06)	(0.23)	(0.07)	(0.08)	(0.04)
Constant	0.23	158.32***	13.20	13.97	0.05	-0.12
	(0.39)	(35.75)	(11.39)	(11.66)	(0.07)	(0.16)
Adjusted R^2	0.13	0.11	0.10	0.09	0.06	0.65
N	129	123	124	124	124	129
Treatment (primary education)	0.12	4.81	0.91	0.51	-0.10	-0.02
	(0.11)	(14.95)	(3.72)	(2.67)	(0.06)	(0.07)
Age 20-24	-0.06	-18.71	-2.16	-1.79	0.04	-0.09
	(0.06)	(28.95)	(3.23)	(2.50)	(0.04)	(0.08)
Age 25-29	0.01	-13.36	-1.15	1.47	0.18**	-0.15*
	(0.14)	(35.75)	(3.13)	(2.78)	(0.07)	(0.08)
Age 30-34	-0.14	4.64	7.35	10.92	0.18*	-0.18**
	(0.08)	(36.52)	(7.66)	(7.83)	(0.09)	(0.07)
Age 35-39	-0.33*	-47.56	-5.93	-0.29	0.26*	-0.14
	(0.17)	(43.18)	(3.90)	(4.16)	(0.13)	(0.14)
Age 40-49	0.06	-74.00	-5.08**	-2.78	0.03	-0.16*
	(0.11)	(60.26)	(2.25)	(2.10)	(0.04)	(0.08)
Female	-0.27*	5.19	-13.30*	-7.70	-0.05	-0.14*
	(0.14)	(29.30)	(7.21)	(7.58)	(0.07)	(0.07)
Never married	-0.15	-38.95*	-1.84	1.67	-0.05	-0.03
	(0.09)	(21.06)	(3.77)	(3.41)	(0.05)	(0.08)
Brahmin and Chhetri	0.04	-22.77	8.66	9.51	0.08	0.01
	(0.08)	(17.04)	(9.47)	(9.66)	(0.06)	(0.10)
Prior training participation	-0.09	-21.32	-10.85	-7.44	0.23**	-0.08
	(0.18)	(24.34)	(10.37)	(11.30)	(0.10)	(0.05)
Y_{t-1}	0.17*	0.08	0.04	0.03	0.07	0.24
	(0.09)	(0.10)	(0.20)	(0.21)	(0.12)	(0.25)
Constant	0.80***	193.36***	11.74***	0.85	0.14**	0.31**
	(0.11)	(31.52)	(3.18)	(3.24)	(0.06)	(0.11)
Adjusted R^2	0.23	0.15	0.14	0.16	0.43	0.14
N	203	184	193	193	193	203
Treatment (secondary education)	0.24	-0.68	-0.67	0.70	0.05	-0.04
	(0.20)	(21.76)	(3.88)	(1.93)	(0.09)	(0.18)
Age 20-24	0.04	-9.39	2.42	-1.05	0.01	0.05
	(0.13)	(37.42)	(3.03)	(0.91)	(0.05)	(0.13)
Age 25-29	-0.08	4.58	3.43	2.07	0.09	-0.10
	(0.31)	(56.08)	(5.85)	(4.59)	(0.11)	(0.27)
Age 30-34	-0.06	24.34	13.26	9.04	0.23	-0.02
	(0.17)	(53.64)	(11.22)	(11.11)	(0.17)	(0.27)
Age 35-39	-0.05	27.45	6.09	5.16	0.00	-0.11
	(0.35)	(74.52)	(6.65)	(6.07)	(0.22)	(0.22)
Age 40-49						
Female	-0.14	11.92	1.25	3.00	-0.01	-0.13
	(0.10)	(36.57)	(4.07)	(3.87)	(0.11)	(0.16)
Never married	-0.11	-7.26	-0.09	-0.79	-0.06	0.10
	(0.12)	(30.49)	(2.61)	(2.00)	(0.07)	(0.18)
Brahmin and Chhetri	-0.12**	2.46	0.99	3.91	0.06	-0.03
	(0.05)	(15.19)	(2.51)	(3.12)	(0.05)	(0.13)
Prior training participation	0.14	-9.32	-4.37	-4.49*	-0.18	0.10
	(0.22)	(23.21)	(4.37)	(2.45)	(0.11)	(0.22)
Y_{t-1}	0.11	0.25	-0.01	-0.06	0.12	-0.62***
	(0.10)	(0.19)	(0.09)	(0.05)	(0.15)	(0.05)
Constant	0.67**	173.38***	8.57	1.65	0.13*	0.29
	(0.27)	(43.18)	(5.97)	(3.59)	(0.07)	(0.24)
Adjusted R^2	0.09	0.04	0.17	0.40	0.15	0.16
N	129	112	125	125	125	129

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * p <0.10, ** p <0.05, *** p <0.01.

Table B.4: **ITT effect on TT group by educational groups**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Treatment (no education)	0.30	36.62	3.51*	0.80	0.01	0.12
	(0.31)	(39.20)	(1.69)	(0.98)	(0.07)	(0.10)
Age 20-24	-0.27	-18.70	0.15	1.22	0.03	0.09
	(0.22)	(35.72)	(2.87)	(2.14)	(0.10)	(0.11)
Age 25-29	-0.23	-30.12	1.25	1.33	-0.07	0.03
	(0.17)	(35.50)	(3.20)	(2.19)	(0.05)	(0.07)
Age 30-34	-0.23	-9.67	-2.25	-0.20	-0.02	0.06
	(0.26)	(34.91)	(1.87)	(1.27)	(0.05)	(0.08)
Age 35-39	-0.12	18.76	-3.90	-3.08	0.05	-0.02
	(0.20)	(39.37)	(3.69)	(3.10)	(0.05)	(0.06)
Age 40-49	0.02	76.74**	0.31	-0.51	-0.11***	0.26
	(0.14)	(33.82)	(1.86)	(1.26)	(0.04)	(0.27)
Female	-0.24*	-0.70	-14.51	-13.13	-0.25*	-0.05
	(0.12)	(27.59)	(13.03)	(12.46)	(0.14)	(0.09)
Never married	-0.03	8.16	0.94	0.80	-0.05	-0.04
	(0.17)	(36.81)	(2.71)	(1.82)	(0.08)	(0.05)
Brahmin and Chhetri	0.12	25.76	1.99	-1.50	-0.03	-0.02
	(0.10)	(27.04)	(2.19)	(1.74)	(0.17)	(0.02)
Prior training participation	0.01	-18.13	0.88	1.06	-0.16***	0.05
	(0.26)	(40.48)	(3.14)	(2.80)	(0.06)	(0.08)
Y_{t-1}	0.10	-0.15	0.10	-0.53*	0.33	-0.20***
	(0.10)	(0.11)	(0.21)	(0.25)	(0.20)	(0.06)
Constant	0.87**	192.74***	14.95	12.43	0.40***	-0.03
	(0.39)	(53.77)	(8.71)	(8.98)	(0.12)	(0.12)
Adjusted R^2	0.16	0.13	0.13	0.16	0.34	-0.03
N	121	115	118	118	118	121
Treatment (primary education)	0.14	35.64	0.65	-2.20**	-0.09	0.06
	(0.15)	(26.80)	(1.99)	(0.90)	(0.13)	(0.07)
Age 20-24	0.06	-23.38	1.83	-0.15	0.02	0.01
	(0.08)	(17.02)	(1.34)	(0.92)	(0.06)	(0.09)
Age 25-29	-0.05	-11.86	-0.43	0.27	0.04	-0.07
	(0.07)	(24.89)	(1.67)	(0.95)	(0.08)	(0.07)
Age 30-34	0.06	7.88	5.83*	3.36	0.08	-0.12
	(0.10)	(37.08)	(3.20)	(1.98)	(0.09)	(0.13)
Age 35-39	-0.08	-26.00	-1.28	-1.36	0.07	0.03
	(0.13)	(27.89)	(2.39)	(2.34)	(0.18)	(0.06)
Age 40-49	-0.32**	-8.02	-6.29**	-4.16*	-0.26*	0.04
	(0.12)	(28.49)	(2.24)	(2.24)	(0.13)	(0.09)
Female	-0.15	-12.09	-1.98	3.26*	0.31**	-0.22***
	(0.10)	(17.18)	(1.88)	(1.88)	(0.13)	(0.07)
Never married	-0.08	-0.86	-1.97	0.28	0.01	-0.05
	(0.10)	(16.45)	(1.37)	(1.11)	(0.08)	(0.05)
Brahmin and Chhetri	-0.05	7.16	0.27	3.01	0.19	-0.13**
	(0.14)	(39.83)	(2.05)	(1.82)	(0.12)	(0.05)
Prior training participation	-0.03	-5.27	1.84	2.19	0.13	-0.09
	(0.04)	(24.11)	(2.16)	(1.69)	(0.08)	(0.06)
Y_{t-1}	0.19**	0.13*	0.04	0.16	0.14	-0.07
	(0.07)	(0.07)	(0.11)	(0.26)	(0.13)	(0.16)
Constant	0.65***	152.16***	7.87***	1.73	0.08	0.29***
	(0.15)	(31.20)	(1.39)	(1.14)	(0.18)	(0.10)
Adjusted R^2	0.14	0.07	0.11	0.19	0.33	0.03
N	220	188	201	201	201	220
Treatment (secondary education)	0.38*	20.02	8.02	2.71	0.20	0.18
	(0.19)	(18.20)	(7.82)	(3.51)	(0.12)	(0.13)
Age 20-24	0.07	4.93	7.92	-0.33	-0.10	0.12
	(0.15)	(23.37)	(5.30)	(2.54)	(0.14)	(0.13)
Age 25-29	0.13	-2.60	6.03	-2.68	-0.04	0.05
	(0.16)	(31.29)	(5.26)	(4.40)	(0.28)	(0.16)
Age 30-34	0.12	-37.66	-7.82	0.14	0.08	-0.22**
	(0.13)	(40.09)	(13.34)	(7.20)	(0.29)	(0.09)
Age 35-39	0.11	35.40	-12.29	4.89	0.47	-0.26**
	(0.16)	(28.81)	(14.81)	(6.08)	(0.30)	(0.11)
Age 40-49	0.19	209.56***	130.85***	141.99***	0.49**	-0.19
	(0.21)	(39.15)	(7.95)	(5.64)	(0.22)	(0.17)
Female	-0.39**	11.60	-15.24***	-1.36	0.02	-0.25**
	(0.15)	(23.43)	(4.43)	(3.46)	(0.11)	(0.09)
Never married	-0.16*	-37.16**	-9.44	-4.72	-0.01	-0.11
	(0.08)	(16.69)	(7.73)	(5.01)	(0.10)	(0.10)
Brahmin and Chhetri	0.04	21.28	9.61	1.55	-0.03	0.06
	(0.09)	(27.15)	(6.54)	(2.82)	(0.09)	(0.13)
Prior training participation	0.16**	9.19	0.32	3.78	0.08	-0.12
	(0.07)	(38.63)	(3.92)	(4.35)	(0.06)	(0.08)
Y_{t-1}	0.01	-0.01	-0.04	0.05	0.14	
	(0.15)	(0.10)	(0.24)	(0.05)	(0.11)	
Constant	0.57***	174.18***	10.28	4.44	0.13	0.13
	(0.15)	(27.63)	(10.73)	(3.13)	(0.15)	(0.19)
Adjusted R^2	0.29	0.25	0.15	0.57	0.16	0.21
N	110	99	108	108	108	110

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.5: **ITT effect on RT group by income groups**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Treatment (low income)	0.24*	19.93	1.51	0.90	0.02	-0.01
	(0.11)	(14.76)	(1.82)	(1.31)	(0.08)	(0.05)
Age 20-24	-0.12	-37.27	-1.20	-0.71	-0.01	-0.06
	(0.08)	(23.51)	(2.04)	(0.71)	(0.04)	(0.06)
Age 25-29	-0.08	-20.62	-3.21	-1.04	0.07	-0.09
	(0.16)	(30.02)	(2.51)	(1.92)	(0.06)	(0.08)
Age 30-34	-0.19	-18.93	1.60	3.67	0.09	-0.10
	(0.13)	(39.97)	(4.27)	(3.84)	(0.08)	(0.09)
Age 35-39	-0.30	-47.91	-6.50*	-2.84	0.03	-0.07
	(0.24)	(38.18)	(3.64)	(3.14)	(0.12)	(0.08)
Age 40-49	0.16	-54.63	-2.53	-0.61	-0.02	-0.06
	(0.16)	(60.98)	(2.58)	(1.00)	(0.06)	(0.06)
Female	-0.27	-13.19	-3.39	2.10	-0.01	-0.16
	(0.16)	(24.78)	(2.55)	(2.75)	(0.12)	(0.10)
Education: primary to below SLC	0.06	-6.95	1.94	2.50	0.16**	-0.00
	(0.07)	(13.09)	(2.28)	(1.90)	(0.06)	(0.04)
Education: SLC and beyond	0.15*	4.02	4.37	6.06	0.21*	0.05
	(0.08)	(17.83)	(4.54)	(4.34)	(0.11)	(0.09)
Never married	-0.10	-28.46	-4.30**	-1.66**	-0.09*	0.05
	(0.12)	(16.32)	(1.75)	(0.73)	(0.04)	(0.07)
Brahmin and Chhetri	-0.06	4.06	2.00	3.15*	0.07	-0.01
	(0.07)	(14.97)	(1.83)	(1.80)	(0.05)	(0.08)
Prior training participation	-0.15	-15.28	-0.46	1.19	0.05	-0.05
	(0.13)	(23.43)	(2.50)	(1.70)	(0.09)	(0.07)
Y_{t-1}	0.16	0.04	3.27	15.39	0.18	0.19
	(0.11)	(0.10)	(6.36)	(14.56)	(0.24)	(0.19)
Constant	0.71***	205.04***	8.41**	-2.55	-0.02	0.29**
	(0.17)	(22.80)	(3.91)	(2.97)	(0.06)	(0.13)
Adjusted R^2	0.14	0.09	0.09	0.19	0.31	0.26
N	281	253	271	271	271	281
Treatment (high income)	0.07	7.21	-0.37	-0.70	-0.01	0.02
	(0.08)	(15.26)	(3.46)	(2.93)	(0.07)	(0.08)
Age 20-24	0.17*	42.90	8.41*	2.30	0.09	0.05
	(0.09)	(27.78)	(4.22)	(2.60)	(0.09)	(0.09)
Age 25-29	0.13	8.80	10.46	9.07	0.25*	-0.09
	(0.14)	(32.84)	(7.38)	(6.78)	(0.13)	(0.11)
Age 30-34	-0.02	31.90	14.73	13.47	0.24**	-0.04
	(0.12)	(29.19)	(10.21)	(9.83)	(0.11)	(0.10)
Age 35-39	0.04	22.14	5.26	4.56	0.32**	-0.14
	(0.13)	(26.08)	(4.09)	(3.54)	(0.13)	(0.12)
Age 40-49	0.20	18.76	7.64	0.25	0.03	-0.04
	(0.14)	(29.82)	(7.00)	(6.69)	(0.10)	(0.11)
Female	-0.18*	-19.55	-19.48	-17.16	-0.08	-0.13*
	(0.09)	(21.60)	(17.50)	(17.50)	(0.10)	(0.06)
Education: primary to below SLC	0.16**	-13.41	5.95	6.02	0.11	-0.10
	(0.07)	(18.45)	(4.00)	(3.53)	(0.12)	(0.11)
Education: SLC and beyond	0.12	-0.64	8.96**	6.10	0.10	-0.07
	(0.13)	(14.64)	(4.18)	(5.12)	(0.16)	(0.13)
Never married	-0.03	-15.43	2.12	1.20	-0.03	0.07
	(0.11)	(17.77)	(3.60)	(3.52)	(0.08)	(0.11)
Brahmin and Chhetri	-0.06	-8.09	8.52	9.99	0.05	-0.02
	(0.08)	(13.74)	(9.46)	(9.41)	(0.07)	(0.09)
Prior training participation	-0.11	-8.98	-8.87	-4.97	0.05	0.08
	(0.15)	(28.31)	(6.09)	(6.90)	(0.11)	(0.08)
Y_{t-1}	0.04	0.08	0.03	-0.06	0.12	-0.03
	(0.11)	(0.10)	(0.11)	(0.16)	(0.12)	(0.20)
Constant	0.64***	160.27***	0.84	-3.51	-0.10	0.23*
	(0.13)	(41.11)	(7.30)	(6.45)	(0.17)	(0.11)
Adjusted R^2	0.12	0.22	0.08	0.13	0.08	0.16
N	180	166	171	171	171	180

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: **ITT effect on TT group by income groups**

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Treatment	0.37*	48.06	2.20	0.04	0.08	0.05
(low income)	(0.19)	(29.39)	(1.67)	(0.81)	(0.09)	(0.06)
Age 20-24	-0.10	-13.77	0.86	-0.27	-0.06	0.07
	(0.07)	(16.45)	(1.36)	(0.95)	(0.07)	(0.06)
Age 25-29	-0.08	-10.57	2.18	-2.25	-0.05	0.05
	(0.08)	(19.22)	(2.12)	(1.77)	(0.08)	(0.05)
Age 30-34	0.03	-3.05	0.73	0.42	0.10	-0.06
	(0.08)	(18.24)	(1.15)	(1.54)	(0.11)	(0.05)
Age 35-39	-0.01	-9.90	-2.87	-3.23	-0.02	0.12
	(0.12)	(30.23)	(2.43)	(2.36)	(0.11)	(0.12)
Age 40-49	0.10	53.10	1.72	-1.79	-0.05	0.17
	(0.12)	(40.59)	(2.34)	(1.80)	(0.07)	(0.15)
Female	-0.24*	-0.64	-6.44**	1.78	0.22**	-0.25***
	(0.12)	(17.17)	(2.35)	(1.34)	(0.10)	(0.06)
Education: primary	-0.10**	-5.84	-2.35	-0.43	0.05	0.01
to below SLC	(0.05)	(15.69)	(1.39)	(1.11)	(0.05)	(0.05)
Education: SLC and	0.01	-5.90	-0.86	-1.03	0.07	0.04
beyond	(0.07)	(26.33)	(2.21)	(1.53)	(0.07)	(0.05)
Never married	-0.12	-18.45	-0.34	-1.59	0.01	-0.06
	(0.09)	(15.13)	(1.18)	(1.81)	(0.07)	(0.07)
Brahmin and Chhetri	0.05	19.69	4.01	2.61	0.06	-0.07
	(0.08)	(18.27)	(3.30)	(1.83)	(0.07)	(0.06)
Prior training	-0.08	4.68	-0.95	2.20	0.02	-0.05
participation	(0.09)	(28.71)	(1.77)	(2.90)	(0.06)	(0.06)
Y_{t-1}	0.04	0.07	-4.46*	-0.11	0.32	-0.26**
	(0.09)	(0.10)	(2.54)	(3.16)	(0.37)	(0.10)
Constant	0.65***	152.49***	8.99***	2.68*	0.00	0.23***
	(0.20)	(37.75)	(2.00)	(1.37)	(0.14)	(0.05)
Adjusted R^2	0.19	0.05	0.20	0.13	0.20	0.14
N	297	262	282	282	282	297
Treatment	0.11	14.29	1.60	-0.34	-0.14	0.07
(high income)	(0.13)	(19.01)	(3.29)	(3.27)	(0.10)	(0.07)
Age 20-24	0.11	0.35	2.15	-2.43	-0.05	-0.01
	(0.11)	(24.77)	(3.15)	(1.68)	(0.09)	(0.09)
Age 25-29	0.08	-19.65	-3.10	3.68	-0.07	-0.14
	(0.13)	(31.32)	(6.24)	(3.73)	(0.07)	(0.11)
Age 30-34	0.01	3.07	-5.43	2.00	-0.03	-0.12
	(0.18)	(40.48)	(6.20)	(2.32)	(0.09)	(0.11)
Age 35-39	0.15	49.33	-2.06	0.17	0.28**	-0.19*
	(0.12)	(38.53)	(4.80)	(3.85)	(0.12)	(0.10)
Age 40-49	0.26*	210.05***	129.27***	148.17***	0.60***	-0.35***
	(0.13)	(35.50)	(6.93)	(3.57)	(0.14)	(0.09)
Female	-0.16**	-13.38	-3.18	-5.19	-0.01	-0.02
	(0.06)	(17.52)	(6.29)	(5.07)	(0.11)	(0.04)
Education: primary	-0.02	4.46	2.07	-1.61	-0.02	0.05
to below SLC	(0.10)	(21.85)	(6.49)	(4.93)	(0.11)	(0.06)
Education: SLC and	-0.18	-16.77	10.47	-2.90	-0.14	0.01
beyond	(0.15)	(26.04)	(12.81)	(5.15)	(0.14)	(0.11)
Never married	-0.01	9.23	-11.77*	-1.62	-0.05	-0.11
	(0.10)	(18.90)	(6.26)	(2.72)	(0.07)	(0.12)
Brahmin and Chhetri	-0.18*	13.71	2.74	1.67	-0.02	-0.01
	(0.09)	(32.40)	(4.71)	(5.02)	(0.09)	(0.08)
Prior training	0.26**	-8.59	-2.10	3.33	0.15	-0.15***
participation	(0.10)	(30.07)	(5.43)	(2.63)	(0.11)	(0.05)
Y_{t-1}	0.05	0.00	-0.10	-0.03	0.18*	0.08
	(0.10)	(0.08)	(0.14)	(0.10)	(0.10)	(0.19)
Constant	0.77***	173.02***	13.86**	6.36	0.31	0.20
	(0.20)	(44.20)	(5.30)	(4.23)	(0.18)	(0.13)
Adjusted R^2	0.08	0.23	0.17	0.58	0.44	0.00
N	156	142	147	147	147	156

Note: SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table B.7: **ITT effect by subgroups**
(Model with interactions of treatment with the characteristics we examined for heterogeneity)

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
a. Randomly selected trainees						
Treatment × female	0.08	24.04	1.16	-1.34	0.05	0.05
	(0.16)	(23.11)	(2.64)	(2.01)	(0.12)	(0.06)
Treatment × primary	-0.00	-17.70	2.77	4.67	-0.08	-0.07
	(0.15)	(27.80)	(6.41)	(5.54)	(0.07)	(0.07)
Treatment × secondary	0.08	-19.67	2.27	5.94	0.11	-0.10
	(0.21)	(29.60)	(6.13)	(4.98)	(0.09)	(0.14)
Treatment × high income	0.02	7.36	1.67	1.04	-0.03	-0.01
	(0.06)	(9.56)	(1.30)	(1.24)	(0.05)	(0.04)
b. Trainer selected trainees						
Treatment × female	0.02	33.45	0.86	0.91	-0.00	0.05
	(0.23)	(31.72)	(5.82)	(5.44)	(0.13)	(0.09)
Treatment × primary	-0.14	11.67	-0.09	-0.66	-0.13	0.01
	(0.16)	(26.77)	(3.58)	(2.69)	(0.09)	(0.08)
Treatment × secondary	-0.15	-19.39	4.07	3.11	-0.02	-0.16*
	(0.20)	(29.47)	(6.28)	(3.82)	(0.09)	(0.09)
Treatment × high income	0.03	-5.29	2.82*	1.27	-0.03	-0.03
	(0.06)	(12.39)	(1.62)	(1.24)	(0.05)	(0.06)

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34, 35-39, and 40-49 years), gender, education (below primary, primary and secondary), marital status, caste, prior experience of vocational training and the value of the outcome variable at the baseline as well as district fixed effects. SEs are clustered at the district level. The *s indicate the *p-values* from the *t-tests* of a null effect against a two-sided alternative: * p <0.10, ** p <0.05, *** p <0.01.

Table B.8: **Estimated OLS coefficients and the corresponding ATE, ATT and ATU**

	OLS (1)	ATE (2)	ATT (3)	ATU (4)
a. Randomly selected trainees				
Gainfully employed	0.18	0.19	0.20	0.17
Monthly hours worked	14.42	17.17	18.40	13.39
Monthly own income	0.27	0.78	1.03	0.02
Income working for oneself	-0.58	-0.43	-0.35	-0.65
Owns business	-0.00	-0.02	-0.03	0.00
International Migration	0.00	0.02	0.03	-0.00
b. Trainer selected trainees				
Gainfully employed	0.28	0.28	0.27	0.32
Monthly hours worked	36.93	36.83	37.69	34.30
Monthly own income	2.55	2.09	4.65	-5.16
Income working for oneself	-0.24	-0.64	1.72	-7.31
Owns business	0.01	-0.00	0.08	-0.22
International Migration	0.07	0.07	0.07	0.05

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34, 35-39, and 40-49 years), gender, education (below primary, primary and secondary), marital status, caste, prior experience of vocational training and the value of the outcome variable at the baseline as well as district fixed effects. User written Stata command *hettreatreg* has been used for the tests. SEs are clustered at the district level. ATE stands for average treatment effect, ATT stands for average treatment effect on the treated and ATU stands for average treatment effect on the untreated. For details of the estimation methodology, see Słoczyński (2022).