I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16705

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

Shyamal Chowdhury
Syed Hasan
Uttam Sharma

JANUARY 2024

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16705

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

Shyamal Chowdhury
University of Sydney and IZA

Syed Hasan
Massey University

Uttam Sharma
Institute for Social and Environmental Research Nepal (ISER-N) and University of Michigan

JANUARY 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

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

Based on a randomized controlled trial conducted on extremely poor youths in Nepal, we report the impact of a vocational training program that offered long-duration training combined with incentives for trainers tied to trainees' success. Furthermore, to mimic the practices in the field, a component of the program allowed trainers to select trainees from eligible applicants. For the trainees that were randomly selected, after nine months of 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. The findings are consistent with the observed pattern of finding a better outcome when the program implementers non-randomly select the treatment groups. Our investigation thus points out that trainee selection can provide a better outcome of vocational training.

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

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

Corresponding author:
Shyamal Chowdhury
School of Economics
University of Sydney
NSW 2006
Australia

E-mail: shyamal.chowdhury@sydney.edu.au

* The IRB approval was obtained from the University of Sydney's Human Research Ethics Committee (Project No.: 2013/980), and the trial was 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 participants in the 2023 Australasian Development Economics Workshop (ADEW) and participants of the University of Michigan Development Economics seminar for commening on an earlier draft.

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; Barrera-Osorio et al., 2023; Doerr, 2022; Hirshleifer et al., 2016; Kluve, 2010; Katz et al., 2022). This article re-examines the effectiveness of such training by providing evidence from a carefully designed intensive and incentive-based program. Evaluating a program that offered vocational training in Nepal in a randomized control trial (RCT) setting, we demonstrate that such training may have a limited effect on the targeted outcomes, even in the short run. We subsequently explore the potential ways to improve the vocational training program design and conclude that trainer selection of trainees can be more effective in this regard.

We study and evaluate 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 its effectiveness. Second, the training incorporated an incentive mechanism in which the training providers received part of the payment only if the trainees had been employed within the first six months after the completion of the training. The mechanism would make the trainees more likely to receive the best possible training.

Third, the length of the training program was three months, reasonable compared to many other similar training programs (e.g., Barrera-Osorio et al., 2023). Shorter training, even when highly effective, can generate low benefits that are difficult to detect statistically.[1] Therefore, this study is more likely to detect the benefits of such training, if any. A fourth advantage was the scope to examine the case when trainers selected the trainees as we convinced the policymakers to retain randomly selected (about) half of the eligible candidates for a trial in which the trainers would select the trainees. We compare the outcome with that of the randomly selected trainees to

[1] A meta-analysis in Card et al. (2018), relying on 200 recent studies find training to have modestly positive effects only in the long-run. Since per participant training costs are usually low, the effects imply high returns compared to those in education.

find whether a better outcome–conditional (on the trainer selection of trainees) average treatment effect (CATE)–can be achieved.

The results in this article highlight that vocational training may generate more benefits when trainers select the trainees, as this group can be better motivated, and trainers may have important information on the unobservable characteristics of the trainees. Previous studies on the effectiveness of training programs in developing countries find both positive (e.g., Maitra and Mani, 2017; Alfonsi et al., 2020; Katz et al., 2022) and null (e.g., Card et al., 2011; Cho et al., 2013; Blattman et al., 2020) effects. Therefore, our study can focus on program design that may provide higher benefit.[2]

Random program assignment can provide a more reliable causal effect of training. However, the non-random assignment is common in vocational training and it is worthwhile to examine the consequence of such selection on the outcome.[3] In fact, while most of the vocational training for unemployed youth was found to be ineffective (see, Heckman et al., 1999; McKenzie, 2017), some studies, mostly non-experimental, found the contrary (e.g. Chakravarty et al., 2019; Van den Berg and Vikström, 2022). For instance, using a regression-discontinuity design, Chakravarty et al. (2019) find vocational training to successfully raise non-farm employment and monthly earnings of Nepalese youth.[4]

We addressed this issue by utilizing a suitable setting that allowed us to investigate whether trainer-selected trainees do better than randomly-selected trainees when the pool of trainees is identical.[5,6] This is particularly interesting because the program uses incentive-based payment

[2]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 were found to persist in the long run (Attanasio et al., 2011, 2017).

[3]Both public and private training providers often resist such random assignment, 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 participants selected by the program implementers and those not selected can be wrongly attributed to the program.

[4]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 and 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).

[5]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.

[6]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).

systems to motivate trainers to select candidates with a higher potential for success. Moreover, the program had reasonable training duration, which was likely to generate larger benefits. Evaluating the training programs in this way allowed us to overcome the low statistical power issue faced by many earlier studies (McKenzie, 2017).[7]

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 were selected randomly to match the number of spots available. In the other part, trainers selected whom they wanted to train, and the rest were left out from getting any training. It means that, in the latter case, the selection of trainees was not random but rather chosen by the trainers from the eligible candidates.

We examined whether vocational training benefited low-income youths in employment, working hours, income, business ownership, and international migration. Our investigation revealed that randomly assigned training participants became 18 percentage points (pp) more likely to be employed, but other outcomes did not change significantly. In contrast, when the trainers selected the training participants, their employment prospects, working hours, and the likelihood of international migration increased by 27 pp, 38 hours, and 7 pp, respectively.[8] 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-value for inference and in the potential presence of treatment heterogeneity.

The estimated employment effect for the trainer-selected group is higher than the corresponding estimates of 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 the monetary benefits are close to the poverty threshold.[9] As a result, this study will likely assist the Government of Nepal and other low-income countries by suggesting how to design vocational training and whom to target for maximising the benefit of the programs. Thus our study further contributes to the strand of impact evaluation literature comparing the outcomes under alternative targeting policies.[10]

[7]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.

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

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

[10]A large number of recent studies, using observable characteristics/features in the data and machine learning technique, evaluated alternative targeting policies (e.g., Blumenstock et al., 2015; Aiken et al., 2022, 2023; Athey

By confirming that training can be more effective when trainers select the trainees, this research makes an important contribution to the literature on vocational training programs in developing countries. As trainers in the program receive full payment only when trainees are employed, they are likely to try harder to train their graduates so that they can find employment. In addition, trainers are likely to select the trainees with a higher likelihood of success (e.g., those with higher motivation and/or capability) as trainers may better understand an applicant's potential for success that is not observed by researchers or policymakers. They may also be motivated positively due to the decision-making power bestowed upon them.[11] By comparing the magnitudes of the impacts with and without the trainer selection of trainees, we can form an idea about the contribution of trainer selection. With the subsequent interview of trainers confirming the hypothesized selection mechanism, we join the literature on the design and effectiveness of training programs.

The rest of the paper proceeds as follows. Section 2 presents program background and describes the research design, including sampling procedures, details about the training, and the timeline of activities. Section 3 presents the empirical method and our data. The section also discusses how the attrition issue can affect our results. Section 4 presents the results, including the conducted robustness checks and examination of heterogeneity in the treatment effect. The policy implications and intervention cost recovery issues are discussed in Section 5. Finally, Section 6 concludes the paper along with a discussion on the possibility 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-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 larger anti-poverty initiatives aimed at

et al., 2023). We, however, evaluated the use (and non-use) of trainers' insights on unobservable trainees characteristics in selecting the training participants on different outcome measures.

[11]Intensive training, on the other hand, makes the effect size larger and thus helps to detect it econometrically.

stimulating microenterprise and employment opportunities for low-income people by providing vocational and entrepreneurship training to about 55,000 trainees per year.[12] They were implemented in 23 of Nepal's 75 districts in early 2014, and eligible applicants took the training free of charge (Employment Fund, 2013).

2.2. The training

The training program evaluated in this study had two important features. Firstly, unlike many vocational training programs in low income countries, the programs were more extensive, with each trainee receiving three months of training. Secondly, 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).

In the training 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 and other costs, the training cost was approximately NRs.40,000 ($\approx$ US$400) per participant for training, with slight variations across training types and providers.[13]

2.3. Research design

Our research relies on a randomized controlled trial (RCT) design. Allocation into treatment and control groups from the eligible applicants who applied for the training involved two steps. From the randomly selected first part in the first stage, based on the available spots, a predetermined number of training participants were chosen randomly in the second stage. The remainder constitutes the control group. This treatment and control group provided our *first set of participants* participants.

In the other part, trainers chose the predetermined number of trainees in the second stage. The selected participants were compared against the control participants of the *first set of participants*.

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

[13]We learned about the program cost through personal communications. For more details about the program, including its costs, see Employment Fund (2013).

Thus, the selected trainees and the randomized control group formed our *second set of participants* participants. We also compared the outcomes of the remaining (trainer left out) participants against the identical control group. They together formed our *third set of participants.* So, our research design estimates the effects of a 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

To participate in the training, applicants had to be aged between 18 and 40 years and not enrolled in formal education at the time of the application. EF followed the Government guidelines and relied on the Training & Employment (T&E) providers to select trainees based on their own guidelines (Employment Fund, 2013). Our study's total number of participants was 1,036, a reasonable size compared to the studies on low income countries listed in the review by McKenzie (2017). This study included the 34 vocational training events organized by EF across Nepal at that time. Each event typically trained around 22 trainees, meaning there were about 748 training spots.

Out of the 1,036 applicants who were selected for this research, the *first set of participants* contains 512 applicants, of which 373 persons were randomly assigned to the treatment group (participating in one of the 17 training events). The remaining 139 individuals constituted the (universal) control group. The motivation for selecting a specific number of trainees was to fulfill the available training capacity. From the remaining 524 persons, 373 training participants were selected by the T&E providers (again to fulfill the training capacity) who participated in one of the 17 training events. These 373 training participants and the 139 randomly selected control participants from the *first set of participants* made the *second set of participants* size 512. The remaining (not selected) 151 applicants and the unique control group made the *third set of participants* size 290 (Table 1).

[Table 1]

2.5. Study Timeline and Data Collection

The baseline survey was conducted from March to early April 2014, before the program was implemented. The training programs started in late April and concluded in early July 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 similar sets of questionnaires.

2.6. Data

To collect data from the study participants, we selected Nielsen, a survey firm with a proven track record, through a competitive bidding process. Nielsen collected baseline and endline information for all study participants.

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, and vi) whether the applicant has migrated overseas. Those outcome variables are commonly employed in the studies on vocational training (e.g., Cho and Honorati, 2014; Blattman and Ralston, 2015). They are considered important to indicate the intervention's effectiveness (McKenzie, 2017), and so are the primary outcome variables in our analysis. Note that we use the level values of all monetary dependent variables since using their logarithmic transformations may artificially show very high treatment effects for certain outcomes with values close to zero at the baseline.

Nielsen also collected information for another set of outcome variables which were 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.

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. To mitigate this issue, we conducted the follow-up survey at least three months after the verification of employment, which was the basis for incentive payment.

The information related to the control variables were age, sex, years of education, marital status, and caste. For the continuous independent variables, we converted them into group dummies to make the estimates consistent, as suggested in J-PAL (2022).

2.7. Balance check and summary statistics

We examined the summary statistics for the control and outcome variables collected in the baseline survey to check whether the treatment and control groups were balanced before the intervention, and thus, the setup remains valid for the unbiased estimation of the treatment effect.

Table 2 provides the means and standard errors (SEs) for all control and outcome variables organized under four categories: i) randomly assigned control group (Column 1), ii) randomly assigned trainee group (Column 2), iii) trainer assigned trainee group (Column 4), and iv) the group not selected by the trainers (Column 6). As the first category is the universal control group in our investigations, while the last three sets have been used as treatment groups in separate analyses, we also present the difference in means between the treatment and control groups (Columns 3, 5 & 7). The table shows that, for all three pairs of treatment and control groups, there are no systematic differences in the control variables in the baseline (Panel a). The $p - value$ from the $F - test$ of joint significance also shows that, as a group, the differences in the control variables between the treatment and control groups are not jointly significant. With a few exceptions, a similar pattern is observed in all three cases regarding the outcome variables (Panel b).

[Table 2]

2.8. Attrition

Of the 1,036 study participants in the baseline, 241 (23.2%) could not be contacted in person in the follow-up survey. Of those, 80 applicants were outside Nepal, and their interviews were taken over the phone using a shorter questionnaire. The remaining 161 participants (15.5%) attrited during the endline, and the attrition rates were slightly higher for the control group than the randomly selected or the trainer selected trainees (Appendix Table A.1).[14]

To investigate the impact of treatment assignment on attrition, we regressed attrition on treatment assignment and the control variables using our main specification (equation (1) in Section 3), as suggested in J-PAL (2022). We found that treatment assignments were negatively associated

[14] 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 hours, 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.

with attrition of study participants for all three treatment groups, although none were significant at the 5 percent level (Appendix Table A.2).

In the next step, for all three study groups, we examined the difference (with regard to the control and outcome variables) between the treatment and control applicants who attrited. Results indicate no systematic differences in the characteristics of treatment and control individuals (Appendix Table A.3, Panel a). The F-test of joint significance also indicates an insignificant difference when we consider all the control variables jointly. Importantly, there were no significant differences with respect to the outcome variables between the treatment and control individuals for all three study groups in the attired sample (Appendix Table A.3, Panel b). This finding is also confirmed by the F-test of joint significance. Nevertheless, to address any concern about the missing values and attrition, we conducted attrition-adjusted tests throughout the analysis to ensure that our estimate of the treatment effect remains valid for policy purposes.

3. Empirical method

With a randomized setting, we used a linear regression model for our investigations, as follows

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

where, for each individual i, y is one of the six outcome variables discussed in Subsection 2.6. $Treatment$ is a binary variable taking the value of 1 if the individual belongs to the treatment group and 0 otherwise. $\boldsymbol{X}$ is a vector of baseline-level controls for individual and household characteristics, including the baseline outcome. φ_d represents district fixed effects (FEs) while ε is a mean-zero error term.

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). Furthermore, as a robustness check, using training assignment as an instrument for actual training participation, we also estimated the interventions' local average treatment effects (LATEs) and discussed them in Subsection 4.5.

In this study, we followed certain norms to improve the quality and consistency of the analysis. First, as suggested in Athey and Imbens (2017); Wooldridge (2021); Abadie et al. (2023), we used robust standard error to account for heteroskedasticity and clustered them at the district level to address the issue that treatment assignment was based on the available training spots in districts. Second, we set a seed and used 1,000 replications for bootstrapping to ensure the replicability of the results. Third, we followed the norm of using a 5 percent significance level for hypotheses testing.

4. Results

We begin our investigation by comparing the mean outcomes in the endline between the treatment and control group participants for all three sets of participants. For the *first set of participants*, in which we compared randomly selected treatment and control participants, the results indicate a statistically (and practically) significant effect on employment but not jointly for all outcomes (Appendix Table A.4). For the *second set of participants*, in which we compared trainer-assigned participants with the randomly selected control group, gainful employment, working hours and business ownership were significantly higher for the treatment group. Furthermore, when we conducted an $F-test$ of joint significance of all the outcome variables, we rejected the null hypothesis of no difference between the treatment and control groups. For the *third set of participants*, in which we compared participants who were not selected by the trainers with the randomly selected control group, income was significantly lower, but we could not reject the null hypothesis that there were no differences in the outcome variables jointly.

In the next step, we estimate the intention to treat effects (ITT) separately for all three sets of participants by running the regression model specified in equation (1) for all outcome variables. Additionally, we conducted robustness checks of our estimates of the treatment effects in each part of our analysis. First, we used regression adjustment (RA) that contrasts the averages of treatment-specific predicted outcomes to estimate the TEs. The method is useful when there is a selection bias in the RCTs that generally produces misleading results (Allcott, 2015; Słoczyński, 2022; Krauss, 2018, 2022). RA can produce the TE estimates that are robust of any potential selection bias.

Second, we employed inverse probability-weighted regression adjustment (IPWRA) that 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 treatment effects.[15] Third, we used augmented inverse-probability weighting (AIPW) with the selection of covariates using a machine learning approach (Lasso method).[16] Fourth, we estimated the Lee bounds and tightened them by adding covariates, as suggested in Lee (2009) and J-PAL (2022).[17] Fifth, we examined the significance of the estimated TEs with the randomization inference method.[18]

4.1. Effect on randomly assigned group

Table 3 presents the results of our *first set of participants*, which examines the effects of the training on the outcome variables when trainees are randomly selected. Panel a presents the results from the estimating equation (1). Column 1 results show that training significantly improves the probability of being gainfully employed by 18 percentage points (pp). The coefficient is large and indicates the success of the intervention to some extent.[19]

[Table 3]

The results remain consistent when we estimate the treatment effect with regression adjustments (Panel b), inverse probability weighting (Panel c), and regression adjusted inverse probability weighting and selecting covariates using Lasso (Panel d). The Lee bounds also confirm a significant effect on the outcome even with very conservative assumptions (Panel e), while the randomization inference test results (Panel f) confirm that employed significance that uses the standard $t-distribution$ is similar to those observed in our data.

[15] 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.

[16] 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. Thus, the AIPW estimates are more likely to employ better outcome models to estimate the TE.

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

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

[19] Among the other variables in the model, the participant's sex and previous employment status are statistically significant, suggesting lower employment prospects for female applicants and higher for participants who are already employed. These findings reflect the commonly observed pattern in the labor market, where females face societal challenges in entering the job market. On the other hand, already employed applicants may have the capacity and network to either continue with their current job or find a new one, making their effect significant in our model.

Training does not seem to have any significant impact on the other outcome variables we considered in this analysis (Columns 2-7). 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 indicates that the benefit of being employed may not be high for the training participants. The treatment has a limited (and insignificant) effect when we include home cultivation in defining gainful employment (Appendix Table A.5).[20]

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 and Ralston, 2015). As a result, they may have the flexibility of reporting their employment status either way.

Therefore, 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 outcome of the extreme poor. These findings are consistent with most previous studies and reflect the fact that without capital, the returns to technical skills could be limited or that designing useful training programs can be challenging (Heckman et al., 1999; Kluve, 2010; Blattman and Ralston, 2015; McKenzie, 2017).[21,22]

4.2. Effect on trainer assigned group

In Table 4, we present the results of the effects of training on all primary outcomes when the trainers select the trainees as opposed to random selection. Panel a presents the results from the estimating equation (1). The estimated TE in Column 1 indicates a 27 pp increase in the probability of gainful

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

[21]Nevertheless, there are studies finding positive effects of vocational training in the short-run (e.g., Maitra and Mani, 2017; Doerr, 2022; Baird et al., 2022; Adhvaryu et al., 2023). Maitra and Mani (2017) find a subsidized vocational education program for women residing in low-income Indian households to increase participants' employment, working hours, 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).

[22]The findings in Balboni et al. (2022) can be 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 households who begin in extreme poverty 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.

employment. The effect is about 50 percent higher than the impact on the randomly selected group and is statistically significant at the 5 percent level against a one-sided alternative. Other significant controls indicate that females are less likely to be employed, while previously employed trainees have a higher potential for employment, as seen earlier. The effect also remains significant when we employ RA (Panel b), IPW (Panel c), AIPW (Panel d), Lee bound (Panel e), and randomization inference (Panel f) in our analysis.

[Table 4]

Training also has a significant impact on the working hours of the treatment group. The estimate indicates that the trainer-assigned trainees worked 38 hours more per week 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 in Panels b-f. Their income also increased by NRs.2,800, but it was not statistically significant (Column 3). Understandably, their monthly income from working for him/herself and business ownership is not affected (Columns 4-5) as they were trained to get employed. However, their international migration also significantly (against a one-sided alternative) increased by 7 pp compared to the control group. The effect on international migration, however, fails to satisfy the robustness checks we considered.

Thus, our analysis indicates some impact on the trainer-selected trainees on employment, working hours, and international migration. Robustness checks with similar outcome variables indicates a like but less significant impact (Appendix Table A.6). So, we conclude that the trainer-chosen trainees generally experience a better outcome than those selected randomly.

We later communicated with some training providers over the phone to know about their selection criteria.[23] They suggested that trainers would primarily look for the likelihood of applicants' taking a full-time job. They observed whether study participants' actions was consistent with their commitment to work. For instance, the trainers embraced if applicants suggested visiting potential employers for job seeking. Similarly, trainers appreciated applicants willing to pay the training fees if dropped out of the program. Referrals from the previous cohort of trainees were also greatly valued. Some training providers give priority to applicants who have family members already working in the same profession. This preference stems from the belief that familial connections can lead to

[23]ISER-N IRB Approval No: A-014/2080/081, Date: August 25, 2023.

improved networking and, thus, increased prospects of employment. In short, they would try to delve deeper to gauge the attitude of the applicants. It may also be due to the selected trainees' comparative advantage in vocational training, as seen earlier in Silliman and Virtanen (2022).

4.3. Effect on the group not selected by trainers

Finally, we examine the impact on the participants who were left out by the trainers. Members of the group also do not take any training like the randomly chosen control group members. They, however, can suffer from negative selection by some unobservable characteristics and thus can experience deteriorated outcomes. Table 5 results show generally negative but no significant changes for the left-out group members. The treatment coefficient (identifying the left-out group members) also remains insignificant when we estimate the changes with the use of RA (Panel b), IPW (Panel c), AIPW (Panel d), Lee bound (Panel e) and randomization inference (Panel f). Further robustness checks with competing outcome variables also find similar results (Appendix Table A.7).

[Table 5]

Thus, the overall results indicate that the left-out group members generally do not experience a deterioration in the outcomes. This can be because the applicants operate in the low-skilled job markets, which only require a little ability, motivation, and networking capacity, and thus, they experience a similar outcome in the job market.

4.4. Comparison of effects among different treatment groups

Until now, we examined the changes separately by study groups to allow for their differences in relation with the control variables. At this stage, we combine the groups together to benefit in statistical significance of the coefficients of interest. We run model (1) on all the study participants with the randomized control group as the reference category and add treatment indicators for all three groups (to separately estimate the effects/changes for the three treatment groups).[24]

Panel a results in Table 6 indicate a similar pattern of impact/changes we observed earlier. Specifically, it provides a significant estimate of the effect on the employment of the randomly

[24] Specifically, we use the model: $y_i = \beta_0 + \beta_1\, Treatment_{1i} + \beta_2\, Treatment_{2i} + \beta_3\, Treatment_{3i} + \boldsymbol{\theta X_i} + \varphi_d + \varepsilon_i$, in which, in addition to the earlier notations, $Treatment_g$, $g = \{1, 2, 3\}$ indicates the treatment group from the g^{th} set of participants.

selected trainee group. The effects were again not statistically significant for any other outcome variables. On the other hand, the trainer-selected trainees experience a positive impact on employment and working hours. Interestingly, for both of the outcome variables, the estimated impacts were higher for the latter group. We see no significant impact on the trainer left out applicants.

[Table 6]

In the previous analyses, we did not compare the observed post-intervention differences between the groups. In Panel b of Table 6, rather than three separate treatment indicators, we added a treatment indicator and its interaction (separately) with the other two study group indicators. The results indicate about 9 pp higher impact of training on the employment of the trainer selected group. The difference was almost statistically significant at the 5 percent level against a one-sided alternative. The group also gains in working hours by 15.2 hours per month. Benefits regarding other outcomes of interest were not statistically significant. On the other hand, the trainer left out group faces insignificant reductions in employment and other outcomes, although the negative changes in working hours and income appear to be large.

To better compare the effect on outcomes for the three groups, Figure 1 below presents the standardized effects for all six outcomes. The figure shows that randomly selected trainee groups gain only with regard to employment, while the trainer-selected group members gain more in employment while also raising their working hours and international migration. The trainer left-out group remains roughly similar to the universal control (randomly selected) group.

[Figure 1]

Our estimated TEs are modest for the randomly selected participants and are broadly consistent with Heckman et al. (1999) who suggested vocational training programs to generate low benefit as they generally target low-quality participants. Our finding that the trainer-selected group experiences a (slightly) superior outcome is intuitive, as trainers may better understand applicants' ability, suitability, and urge for jobs, as we hypothesized. It 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 conducted an RCT in West Africa and found that personal initiative training, but not traditional training, improves outcomes.

At this point, a more appropriate interpretation of our results is worth discussing. The effect of selection and the effect of training may not be additive. In particular, when only better-quality applicants are trained, the outcome will improve through a) the training, b) the quality (including better matching) of participants and c) their interactions. Our comparison of the randomly selected trainees with the control group (Table 3) offers an idea of (a). We also observed no significant effect on the applicants who were not selected by the trainers (Table 5), indicating a likely limited contribution of (b). Thus, the improvement of the trainer-selected trainees over the randomized treatment group is likely due to the interaction effect (c). The overall results thus suggest that, while the selection does not affect the outcomes directly, it does so indirectly through the interaction with training. The indirect effects may stem from, among others, the heterogeneous effect of training with regard to ability, matching, and knowledge–that the trainers can guess about the applicants during the trainee selection process.[25]

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

4.5. Robustness and Heterogeneity

Our investigation relies on many 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 of the treatment variable, adjusted for multiple hypothesis testing, we follow the approach provided in Romano and Wolf (2005a,b, 2016) and Clarke et al. (2020). Table 7 below presents three types of $p-values$ for the treatment effects on the outcome variables for each study group. In the table, Columns 1, 4, and 7 present simple (uncorrelated model) $p-values$, Columns

[25] A foolproof way to identify the causal effect of trainee selection can be achieved through, for example, i) randomly dividing the participants into treatment and control groups and ii) for both of the groups, allowing the trainers to select the trainees—say trainer selected and not selected subgroups—blindly (not knowing which is treatment and which is control group). The impact of training on a person the trainer would have chosen can then be estimated by comparing the trainer-selected treatment and control subgroups.

2, 5, and 8 present the $p-values$ by random permutation respecting strata and clusters, while Columns 3, 6, and 9 present the Romano-Wolf (R-W) multiple hypotheses corrected $p-values$.

[Table 7]

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 randomly selected group, ii) employment and working hours when trainers selected applicants, and iii) no changes for the trainer left-out group. The table also presents Randomization $p-values$ for joint tests of treatment significance, as discussed in Young (2019). In the first two cases, we reject the null hypothesis that training improves none of the outcomes but cannot reject the same for the final group.

Next, we confirm that the treatments are not made significant by $p-hacking$. 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 treatment variable. We generate standardized graphical outputs from regression specifications by individually regressing a dependent variable against all possible combinations of independent variables. The effect curves (histograms of the estimated treatment coefficients) and the t-curves (histograms of the absolute value of the $t-statistics$ of the treatment coefficients) closely match the estimates in our employed model, confirming the validity of our estimates.[26]

McKenzie (2017) concludes that the real impact of vocational training is small and thus difficult to identify due to small sample size. All our studies employ long-duration training and the sample sizes appears reasonable, the design may not have enough power to detect a modest effect on many outcomes. As a result, based on observed standard deviations in the actual outcomes for the control groups, we compute the minimum detectable effect size (MDES) for each outcome for which it would have adequate statistical power. We follow the standard practice and consider 80% power with a two-sided test at 5% significance level (Islam et al., 2021). A true positive impact smaller than the corresponding MDES will have less than an 80% chance of being identified.

[26] See appendix, Figures A.1-A.3, 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 each of the three study 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.

For the randomly assigned group, our estimated effect size is larger than the MDES (in their original units of measurement) for employment only. For the trainer-assigned group, the estimated effects are larger than MDES for employment, working hours, and income. No estimated effects are larger than MDES for the trainer left-out group (Appendix Table A.8). Therefore, failure to identify any effect of training is less likely to be a problem in our case. Thus, the comparison of MDES with our estimated effects generally supports our conclusion of a lower effect of vocational training on the randomly selected trainees but a larger impact on the trainer-assigned trainees.

In all three studies, some applicants assigned to the control group took the training, while some applicants assigned to the treatment group did the opposite. Thus, the ITT estimates are likely to underestimate the true treatment effects. Therefore, we estimated the local average treatment effects of training participation in all the cases we analyzed earlier. In our estimation, we used treatment assignment as an instrument of actual training participation and estimated models for all the outcomes. Our estimation of LATE otherwise follows specification (1) that we used earlier to estimate the ITT effects. As expected, the LATE estimates are slightly higher than the ITT estimates (Appendix Tables A.9-A.11). Importantly, our conclusions about the effectiveness of training for all three studies remain unaffected when we consider the statistical significance of the coefficients of the treatment effects.

Heterogeneity in the treatment effect is commonly observed in empirical studies evaluating vocational training (Blattman and 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), caste (Field et al., 2010) and unobserved ability (Rodríguez et al., 2022). One particular problem in our case is that the OLS estimation of equation (1) is generally inappropriate in the presence of heterogeneity (Słoczyński, 2022). Therefore, we repeated the previous analysis by sex, education, and income subgroups.[27,28]

[27] 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 and Mckenzie, 2023).

[28] It should be noted that discovering and exploiting heterogeneity of the TEs is not a key goal of this research. Thus, the subgroup analysis aims to show whether the 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 subgroups and a larger sample than examining whether the treatment has an effect (Duflo et al., 2007; Chernozhukov et al., 2018).

Table 8 presents the estimated TEs (and their SEs) for the outcome variables by study groups and subgroups defined by sex, education, and income (detailed in Appendix Tables B.1-B.9). Ignoring the statistical significance for now, we usually observe a positive impact on employment for all subgroups of the randomly assigned trainees (Panel a). In addition, female, low-educated, and low-income participants benefit more. For the trainer-assigned trainees, we generally observe a positive impact on employment, working hours, and income for all subgroups (Panel b). Again, female, low-educated, and low-income participants benefit more. The effects are mostly higher for the trainer-selected group for all subgroups. The changes for the trainer left out group are mostly negative but low.

[Table 8]

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 suggest that the success of job-training programs may depend on trainees' expectations, as found in Acevedo et al. (2020).[29] Education is also likely to interact with the training through productivity (positively) and motivation (negatively). For example, Bassanini (2004) find training to have a stronger impact on employment security in the case of low-educated workers. The higher effect on low-income individuals may be due to their motivation and urgency for 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.

We also estimated models with interactions of treatment and characteristics and concluded similarly (Appendix Table B.10). To further confirm that heterogeneity does not invalidate the treatment effect, we use the method provided in Słoczyński (2022). The results largely indicate that the estimated TEs are not much different from ATT or ATE (Appendix Table B.11). Thus, the heterogeneity analysis is largely consistent with the existing literature and provides support to our findings—vocational training may benefit the trainees more when trainers choose the participants.

[29] 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., 2021).

5. Policy relevance and intervention cost recovery

Developing countries, such as Nepal, are continually seeking ways to improve the economic status of their relatively poor population. In this regard, vocational training, which we evaluated here, is an approach to enhance labor productivity and increase employment opportunities.[30] Thus, our findings may have important policy implications in this context. Firstly, we confirm that even intensive vocational training, combined with trainers' incentives linked directly to trainees' employment, only affect their employment prospect. This finding is consistent with many previous studies reporting a null or small effect of vocational training (Heckman et al., 1999; Blattman and Ralston, 2015; McKenzie, 2017).

We also find that when the trainees are selected by the trainers, the impact on employment is higher while their working hours improve. Local trainers may have information on unobservable trainee characteristics, like urge and motivation for a job and income. Together with the incentive-based payment, the unobservable characteristics can be effective in improving the training outcomes like working hours and income. This can also be a story about signaling and realization of quality, as previous studies found people to possess valuable skills unobservable to employers (e.g., Abebe et al., 2021a,b). Therefore, the finding is likely to have significant implications for the future design of vocational training programs—allow trainers to select the training participants. The study thus contribute to the targeting literature and, in spirit, parallel to the causal machine learning literature evaluating treatment effect heterogeneity (e.g., Aiken et al., 2022, 2023; Athey et al., 2023).

A proper cost-benefit analysis framework, however, will compare 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.2,760 for the trainer-selected applicants. Although none of the estimates are statistically significant, the latter one is economically large. With the training cost of NRs.40,000 per trainee, our back-of-the-envelope calculation indicates that the former group will require 12 years while the latter group will need one year and three months to recover the cost of the training. Note that incorporating the domestic interest rate into

[30] Vocational training may have some other beneficial effects on society. 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 important, we focus only on economic outcomes.

the analysis (which is also more appropriate) will further raise the time required to recover the cost of the training, but we have ignored the issue in our calculation for simplicity.

We can also take a return-on-investment approach discussed in detail in McKenzie (2021). De Mel et al. (2008) suggested a five percent monthly return from investing in a microenterprise, and so financing the training cost of NRs.40,000 would earn NRs.2,000 per month. This return appears to be much higher than the income increase of the randomly selected trainees, but the opposite is true for the trainer-selected applicants. Even a one percent monthly return provides a higher monetary benefit than the gain in income for the randomly selected trainees. The cost-benefit analysis thus indicates that vocational training on the randomly selected training will misallocate the resources and so will be counterproductive. However, this will not be the case for the training of the trainer-selected applicants.

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

6. Conclusions

We investigate the impact of intensive and trainers' incentive-based vocational training on applicants' employment, working hours, income, business ownership, and international migration. We find that the training has a limited effect on the outcomes but can generate some benefit when the trainees are selected by the trainers who may have some idea about the unobservable characteristics of the trainees like motivation, knowledge, matching and ability. Our results indicate that positive selection did not directly affect the outcomes but indirectly through their interaction with training. During the trainee selection, the trainers could guess the applicant's quality regarding those characteristics and act accordingly. Interviewing the trainers later confirmed the understanding.

The cost-benefit analysis further indicates that a long time is required for the randomly selected vocational training participants to cover their training cost, but it can be quickly recovered for the trainees selected by the trainers. As this study is one of the most rigorous evaluations of employment training done in Nepal with a randomized control trial research design, the findings can assist Nepal and other low-income countries in designing their policies to promote employment and reduce poverty. It will also attract the interest of the parties involved in the process –training providers, NGOs, government agencies, and international donors.

Thus, it is worth discussing the possibility of scaling up the training programs. We followed List (2022)'s criteria and recommendations to do so. First, for the trainer-selected participants, as we have chosen a large proportion of applicants for training, it's worth exploring whether we can achieve better outcomes with a lower proportion of training participants. Thus, we will need more evidence before scaling up. Second, our study satisfied the sample representativeness criteria as we have chosen applicants randomly for the study. Third, we ensured the representativeness of situations in our study as our participants were selected from the entire interested pool of applicants.

Fourth, the likely spillovers (network effects) and general equilibrium effects of scaling up. While we observed no effect of the treatment on other family members' income, the spillover effects are likely to be positive as long as there is no crowding out. We additionally expect that scaling up will have a positive general equilibrium effect by reducing poverty and vulnerability in the region. Finally, we assessed whether there were any diseconomies of scale associated with our interventions. Since the training mechanism can involve training of the potential trainees who then deliver the training, the interventions are unlikely to suffer from diseconomies of scale.

However, our setup to estimate the causal effect of trainee selection could be improved further by randomly dividing the participants into treatment and control groups in the first stage, while in the second stage, allowing the trainers to select the trainees blindly for both of the groups. The impact of training then could be more precisely estimated by comparing the two trainer-selected subgroups, one each from the treatment and control groups. Therefore, we suggest more investigation on the impact of vocational training, with a design of trainer-selection of trainees—like the one mentioned above—before we recommend scaling up the program.

References

Abadie, A., Athey, S., Imbens, G. W., and 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., and 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., and 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., and Martinez, S. (2020). How job training made women better off and men worse off. *Labour Economics*, 65(2020):101824.

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

Aiken, E., Bellue, S., Karlan, D., Udry, C., and 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., and 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., and 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.

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

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

Athey, S., Keleher, N., and 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 address: https://arxiv.org/pdf/2310.08672.pdf, [Viewed at: 22 October, 2023].

Attanasio, O., Guarín, A., Medina, C., and 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., and 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., and 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., and Heil, A. (2022). Why do people stay poor? *Quarterly Journal of Economics*, 137(2):785–844.

Banerjee, A., Duflo, E., and 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., and 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., and 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., and 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., and 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. and 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.

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

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

Brodeur, A., Cook, N., and 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., and 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., and 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., and 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., and 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. and 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., and 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., Imbert, C., and Rathelot, R. (2021). Can information about jobs improve the effectiveness of vocational training? Experimental evidence from India. IZA Discussion Paper 14427, IZA – Institute of Labor Economics, Bonn, Germany.

Chernozhukov, V., Demirer, M., Duflo, E., and 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. and Honorati, M. (2014). Entrepreneurship programs in developing countries: A meta regression analysis. *Labour Economics*, 28:110–130.

Cho, Y., Kalomba, D., Mobarak, A. M., and 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., and 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., and Zamora, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Quarterly Journal of Economics*, 128(2):531–580.

De Mel, S., McKenzie, D., and 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., and Kremer, M. (2007). Using randomization in development economics research: A toolkit. In Schultz, T. P. and Strauss, J. A., editors, *Handbook of Development Economics*, volume 4, pages 3895–3962. Elsevier.

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

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

Heckman, J. J., LaLonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, pages 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., and 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., and 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. Technical report, The Abdul Latif Jameel Poverty Action Lab (J-PAL). Web address: https://www.povertyactionlab.org/resource/data-analysis, [Viewed at: 18 April, 2022].

Janzen, S. A., Magnan, N., Mullally, C. C., and 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., and 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., and Kaghoma, C. K. (2020). Internal migration and youth entrepreneurship in the Democratic Republic of the Congo. *Review of Development Economics*, 24(3):790–814.

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labour Economics*, 17(6):904–918.

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

Krauss, A. (2022). Assessing the overall validity of randomised controlled trials. *International Studies in the Philosophy of Science*, pages 1–24. forthcoming.

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., Shaikh, A. M., and Vayalinkal, A. (2023). Multiple testing with covariate adjustment in experimental economics. *Journal of Applied Econometrics*, forthcoming.

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

Maitra, P. and 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., and 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. and 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., and Urzúa, S. S. (2022). Dynamic treatment effects of job training. *Journal of Applied Econometrics*, 37(2):242–269.

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

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

Silliman, M. and 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. and 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, 2nd 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. Randomly selected control group	139
b. Randomly selected trainee group	373
c. Trainer selected trainee group	373
d. Applicants left out by trainers	151
Total program participants	1,036

Note: (a+b) makes our *first set of participants*; (a+c) makes our *second set of participants*; while (a+d) makes our *third set of participants*. Total program participants is given by (a+b+c+d).

Table 2: **Balance during baseline by treatment assignment type**

	Randomly selected control group	Randomly selected treatment group		Trainer selected trainee group		Applicants left out by trainers	
	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	0.02	0.40	0.09*
	(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	0.20	-0.03
	(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)	(0.04)	(0.03)	(0.04)
Age 35-49	0.09	0.10	0.01	0.10	0.02	0.07	-0.02
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Female	0.34	0.36	0.02	0.41	0.07	0.26	-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	0.23	-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**	0.50	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	0.35	-0.06	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)	-	-	0.69	-	0.25	-	0.77
Observations	[139]	[373]	[512]	[373]	[512]	[151]	[290]
b. Outcome variables							
Gainfully employed	0.36	0.29	-0.06	0.26	-0.10**	0.32	-0.04
	(0.04)	(0.02)	(0.05)	(0.02)	(0.04)	(0.04)	(0.06)
Monthly hours worked	117.22	119.38	2.16	119.91	2.69	103.44	-13.78
	(8.26)	(5.08)	(9.73)	(5.70)	(10.61)	(9.17)	(12.42)
Monthly own income	3.43	2.54	-0.90	2.22	-1.21**	1.75	-1.68**
	(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	0.64	-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	0.02
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
F-test (p-value)	-	-	0.57	-	0.27	-	0.37
Observations	[139]	[373]	[512]	[373]	[512]	[151]	[290]

Note: Means are reported; SEs are in the parentheses. The columns labelled "Difference" report the mean gaps (and their SEs) between the specific treatment group and the randomly selected control 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 treatment on all the outcome variables in the groups and then tests the null hypothesis that all the slope coefficients are zero. The specific control variables related to the applicants were, age in years (that we categorized as 15-19, 20-24, 25-29, 30-34, and 35-49 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 Certificate (SLC) is 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 collected information on whether they had participated in vocational or skill training earlier. Monetary variables are in thousand Nepalese Rupees. The definition of the variable "Gainfully employed" excludes home cultivation, a proxy for subsistence farming. Prior training participation indicates whether they ever participated in a vocational or skill training.

Table 3: **ITT effect on randomly selected trainees**

	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**	14.41	0.27	-0.58	-0.00	0.00
	(0.08)	(10.81)	(1.43)	(0.90)	(0.05)	(0.04)
b. With regression adjustment						
Treatment	0.19***	14.02	1.01	0.05	-0.02	0.02
	(0.05)	(9.60)	(1.50)	(1.31)	(0.04)	(0.04)
c. With inverse probability weighting						
Treatment	0.19***	14.02	1.01	0.05	-0.02	0.02
	(0.05)	(9.60)	(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***	14.41	0.27	-0.58	-0.00	0.00
Endline control mean	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 and 35-49 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. 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 TE, *"teffects ipwra"* to estimate inverse probability weighted regression adjusted TE and *"telasso"* to estimate inverse-probability weighted TE 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 TE 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 selected trainees**

	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	0.27*	37.80*	2.76	-0.02	0.01	0.07*
	(0.14)	(19.10)	(2.54)	(1.24)	(0.09)	(0.04)
b. With regression adjustment						
Treatment	0.21***	17.56*	1.52	0.23	0.07*	0.01
	(0.05)	(9.96)	(1.59)	(1.15)	(0.04)	(0.04)
c. With inverse probability weighting						
Treatment	0.21***	17.56*	1.52	0.23	0.07*	0.01
	(0.05)	(9.96)	(1.59)	(1.15)	(0.04)	(0.04)
d. With selection of covariates using Lasso						
Treatment	0.19***	17.35*	1.24	0.16	0.09**	0.00
	(0.05)	(9.93)	(1.76)	(1.39)	(0.04)	(0.04)
e. Lee bounds						
lower	0.18***	7.86	-0.65	-0.58	0.04	-0.07
	(0.05)	(12.82)	(2.15)	(1.82)	(0.06)	(0.06)
upper	0.28***	25.58**	2.23	0.40	0.09**	0.04
	(0.07)	(12.72)	(2.01)	(1.45)	(0.04)	(0.04)
f. With randomization inference						
Treatment	0.27***	37.80***	2.76	-0.02	0.01	0.07*
Endline control mean	0.61	171.07	8.46	3.07	0.14	0.15
N	452	403	428	428	428	452

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34 and 35-49 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. 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 TE, *"teffects ipwra"* to estimate inverse probability weighted regression adjusted TE and *"telasso"* to estimate inverse-probability weighted TE 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 TE 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 5: **ITT effect on the applicants left out by trainers**

	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	-0.03	-4.59	-1.41	-0.60	-0.04	0.09
	(0.09)	(20.55)	(1.56)	(0.67)	(0.04)	(0.06)
b. With regression adjustment						
Treatment	-0.08	-18.79	-3.40**	-1.32	-0.01	0.02
	(0.06)	(12.52)	(1.65)	(1.36)	(0.04)	(0.05)
c. With inverse probability weighting						
Treatment	-0.08	-18.79	-3.40**	-1.32	-0.01	0.02
	(0.06)	(12.52)	(1.65)	(1.36)	(0.04)	(0.05)
d. With selection of covariates using Lasso						
Treatment	-0.07	-17.94	-3.59**	-1.58	-0.02	0.05
	(0.06)	(12.45)	(1.60)	(1.31)	(0.04)	(0.05)
e. Lee bounds						
lower	-0.11	-23.53	-4.50**	-1.76	-0.02	-0.00
	(0.07)	(17.70)	(1.96)	(1.48)	(0.05)	(0.06)
upper	-0.08	-14.82	-2.90	-0.24	-0.00	0.03
	(0.08)	(18.89)	(2.61)	(1.95)	(0.07)	(0.05)
f. With randomization inference						
Treatment	-0.03	-4.59	-1.41	-0.60	-0.04	0.09*
Endline control mean	0.61	171.07	8.46	3.07	0.14	0.15
N	240	209	229	229	229	240

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34 and 35-49 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. 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 TE, *"teffects ipwra"* to estimate inverse probability weighted regression adjusted TE and *"telasso"* to estimate inverse-probability weighted TE 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 TE 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 6: **ITT effect on different treatment groups**

	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 all three treatment groups						
Randomly selected	0.17**	13.71	-0.12	-0.80	-0.00	0.01
treatment group	(0.08)	(10.18)	(1.24)	(0.86)	(0.05)	(0.04)
Trainer selected	0.27***	28.90**	3.43	1.50	0.07	0.05
trainee group	(0.09)	(12.51)	(2.74)	(1.93)	(0.09)	(0.06)
Applicants left out	-0.03	-6.46	-2.00	-0.37	-0.03	0.06
by trainers	(0.07)	(15.47)	(1.46)	(1.01)	(0.06)	(0.08)
b. With interaction dummies for different types of Treatment						
Treatment	0.17**	13.71	-0.12	-0.80	-0.00	0.01
	(0.08)	(10.18)	(1.24)	(0.86)	(0.05)	(0.04)
Treatment × trainer	0.09	15.19*	3.55	2.29	0.07	0.04
selected trainees	(0.05)	(7.95)	(2.72)	(2.37)	(0.07)	(0.07)
Treatment × applicants	-0.20**	-20.17	-1.88	0.43	-0.02	0.05
left-out by trainers	(0.09)	(17.11)	(1.46)	(1.50)	(0.06)	(0.09)
N	925	825	875	875	875	925

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34 and 35-49 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 7: **The Romano−Wolf (R−W) multiple hypothesis corrected p−values for treatment**
(With district fixed effects)

	Randomly selected treatment group			Trainer selected trainee group			Applicants left out by trainers		
Outcome Variable	Model (1)	Resample (2)	R-W (3)	Model (4)	Resample (5)	R-W (6)	Model (7)	Resample (8)	R-W (9)
Gainfully employed	0.00	0.00	0.00	0.00	0.00	0.00	0.74	0.71	0.95
Monthly hours worked	0.12	0.14	0.38	0.01	0.01	0.02	0.79	0.79	0.95
Monthly own income	0.88	0.86	1.00	0.34	0.11	0.45	0.51	0.29	0.91
Income working for oneself	0.73	0.70	0.99	0.99	0.99	0.99	0.73	0.56	0.95
Owns business	0.91	0.93	1.00	0.81	0.80	0.93	0.51	0.53	0.91
International Migration	0.94	0.95	1.00	0.20	0.13	0.34	0.18	0.16	0.52
Treatment p-value									
For joint tests		0.00			0.00			0.65	

Note: The reported p-values refer to H_0: β_1=0 against H_1: β_1>0. The $p - values$ in columns 1, 4, and 7 are generated from simple (uncorrelated) model; the p-values in columns 2, 5, and 8 are derived from models that randomly resamples respecting strata and clusters and; the p-values in columns 3, 6, and 9 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 relies on the methodology outlined in Young (2019).

Table 8: **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. Randomly selected trainees						
Male	0.14*	3.44	0.77	0.73	-0.02	-0.02
	(0.07)	(14.70)	(2.40)	(1.68)	(0.05)	(0.06)
Female	0.21	23.68	1.85	-0.46	0.02	0.01
	(0.17)	(19.66)	(1.52)	(0.90)	(0.10)	(0.02)
No education	0.15	28.03	0.25	-2.91	-0.00	0.06*
	(0.15)	(21.12)	(3.80)	(3.74)	(0.05)	(0.03)
Primary education	0.12	4.31	0.92	0.46	-0.11	-0.02
	(0.11)	(14.99)	(3.72)	(2.68)	(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*	20.02	1.45	0.87	0.02	-0.01
	(0.11)	(15.07)	(1.81)	(1.28)	(0.08)	(0.05)
High income	0.07	7.14	-0.32	-0.79	-0.02	0.02
	(0.08)	(15.02)	(3.44)	(2.83)	(0.07)	(0.08)
b. Trainer selected trainees						
Male	0.20	28.77	4.37	1.24	-0.10	0.06
	(0.12)	(17.44)	(4.10)	(2.53)	(0.12)	(0.07)
Female	0.29	58.56	1.86	0.39	0.08	0.08
	(0.35)	(42.68)	(1.87)	(1.24)	(0.05)	(0.05)
No education	0.30	39.15	3.64**	0.78	0.00	0.13
	(0.31)	(37.98)	(1.70)	(0.95)	(0.07)	(0.10)
Primary education	0.14	35.95	0.58	-2.24**	-0.09	0.06
	(0.15)	(27.10)	(2.03)	(0.93)	(0.13)	(0.07)
Secondary education	0.39*	17.02	8.22	2.86	0.20	0.18
	(0.19)	(17.69)	(8.38)	(4.53)	(0.11)	(0.12)
Low income	0.37*	49.47	2.30	0.07	0.08	0.05
	(0.19)	(29.93)	(1.72)	(0.81)	(0.09)	(0.06)
High income	0.11	13.00	2.11	0.21	-0.14	0.07
	(0.13)	(18.29)	(3.60)	(4.03)	(0.10)	(0.07)
c. Applicants left out by trainers						
Male	-0.01	0.19	-0.65	0.14	-0.08	0.13
	(0.13)	(42.64)	(2.58)	(1.54)	(0.08)	(0.12)
Female	-0.03	-25.23	-1.32	-0.54	-0.03	0.00
	(0.08)	(23.49)	(1.61)	(1.45)	(0.10)	(.)
No education	0.00	-16.71	-2.69	-0.63	0.01	0.15
	(0.09)	(19.55)	(2.86)	(2.48)	(0.05)	(0.10)
Primary education	-0.11	3.37	-0.65	-1.13	-0.17**	0.09
	(0.11)	(38.53)	(2.81)	(1.57)	(0.07)	(0.10)
Secondary education	0.15	-84.91	0.10	-1.68	-0.10	0.27**
	(0.32)	(58.22)	(6.03)	(3.47)	(0.18)	(0.10)
Low income	-0.07	-14.73	-2.04	-0.55	-0.08	0.05
	(0.11)	(25.48)	(2.03)	(1.14)	(0.06)	(0.06)
High income	-0.09	4.10	-3.14	-1.97	-0.04	0.14
	(0.09)	(43.86)	(3.45)	(3.07)	(0.07)	(0.09)

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34 and 35-49 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. Study participants reporting nil income in 2014 were considered as low-income.

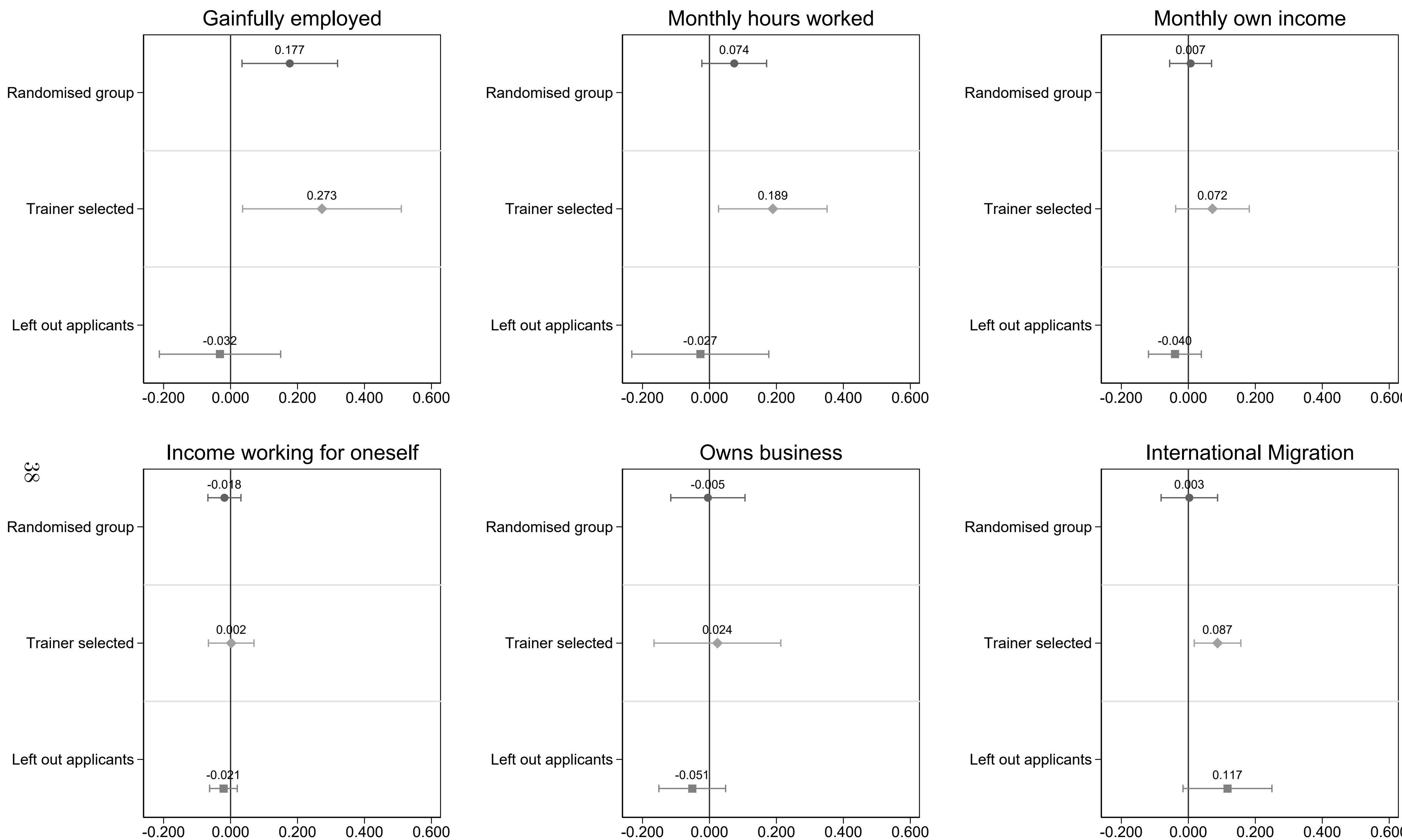

Figure 1: **Group wise standardized effect size (with 95% CI against a one-sided alternative) on selected outcomes**

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. Randomly selected trainees			
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 selected trainees			
Contact in person	100(71.9%)	294(78.8%)	394(77.0%)
Contact over phone	12(8.6%)	22(5.9%)	34(6.6%)
No contact	27(19.4%)	57(15.3%)	84(16.4%)
Total	139(100%)	373(100%)	512(100%)
c. Applicants left out by trainers			
Contact in person	100(71.9%)	103(68.2%)	203(70.0%)
Contact over phone	12(8.6%)	14(9.3%)	26(9.0%)
No contact	27(19.4%)	34(22.5%)	61(21.0%)
Total	139(100%)	151(100%)	290(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 treatment assignment: OLS with the main specification**

	Randomly selected group (1)	Trainer selected trainee group (2)	Applicants left out by trainers (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 35-49	-0.04	0.01	0.04
	(0.07)	(0.08)	(0.14)
Female	-0.13***	-0.07	-0.12*
	(0.03)	(0.05)	(0.06)
Education: primary to below SLC	-0.00	0.01	0.05
	(0.05)	(0.04)	(0.08)
Education: SLC and beyond	-0.04	-0.05	-0.04
	(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	-0.07	-0.12***	-0.21***
	(0.06)	(0.04)	(0.06)
Constant	0.28***	0.28**	0.28**
	(0.07)	(0.12)	(0.10)
Adjusted R^2	0.03	0.06	0.07
N	512	512	290

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**

	Randomly selected control group	Randomly selected treatment group		Trainer selected trainee group		Applicants left out by trainers	
	Mean	Mean	Difference	Mean	Difference	Mean	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(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 35-49	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)
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	0.56	0.49	-0.07	0.60	0.04	0.62	0.06
below SLC	(0.10)	(0.08)	(0.12)	(0.07)	(0.12)	(0.08)	(0.13)
Education: SLC and	0.26	0.35	0.09	0.26	0.00	0.32	0.06
beyond	(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. The value displayed for t-tests are the differences in the means across the groups. 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 assignment type and their differences**

	Randomly selected control group	Randomly selected treatment group		Trainer selected trainee group		Applicants left out by trainers	
	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***	0.55	-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	188.42	17.35*	153.12	-17.94
	(8.63)	(4.69)	(9.58)	(4.99)	(9.90)	(9.05)	(12.52)
Monthly own income	8.46	9.48	1.02	9.77	1.31	4.87	-3.59**
	(1.41)	(1.01)	(1.93)	(1.12)	(2.06)	(0.77)	(1.59)
Income working for oneself	3.07	3.23	0.16	3.39	0.32	1.49	-1.58
	(1.21)	(0.96)	(1.80)	(0.71)	(1.39)	(0.52)	(1.30)
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	-	0.23
Observations	[139]	[373]	[419]	[373]	[403]	[151]	[209]

Note: Stars 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 randomly selected trainees**
(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	1.08
	(0.07)	(0.36)	(0.03)	(0.01)	(1.42)
Age 20-24	0.10	-0.43	0.03	-0.00	-0.57
	(0.07)	(0.73)	(0.07)	(0.01)	(1.05)
Age 25-29	0.14	-0.40	0.02	0.01	-1.05
	(0.08)	(0.88)	(0.08)	(0.01)	(2.58)
Age 30-34	0.14*	-0.06	0.04	-0.00	1.46
	(0.07)	(0.97)	(0.08)	(0.00)	(2.78)
Age 35-49	0.16	-0.79	0.03	0.02	2.09
	(0.09)	(0.93)	(0.08)	(0.02)	(3.11)
Female	-0.28***	-0.50	-0.08	0.02	5.43
	(0.07)	(0.67)	(0.06)	(0.02)	(3.15)
Education: primary	0.09**	-0.36	0.02	0.01	4.68*
to below SLC	(0.03)	(0.26)	(0.02)	(0.01)	(2.56)
Education: SLC and	0.02	0.23	0.00	0.02	10.14**
beyond	(0.04)	(0.40)	(0.06)	(0.02)	(3.83)
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	2.16
	(0.03)	(0.31)	(0.05)	(0.01)	(3.06)
Prior training	-0.05	-0.53	0.06	-0.01	-3.38
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	-1.15
	(0.11)	(0.76)	(0.09)	(0.01)	(3.59)
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***	0.47	0.08***	-0.01	1.20
	(0.04)	(0.32)	(0.02)	(0.01)	(1.26)
N	461	419	398	442	442
c. With inverse probability weighting					
Treatment	0.10***	0.47	0.08***	-0.01	1.20
	(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	1.08
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 selected trainees**
(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	1.26*	0.05	0.00	-2.09
	(0.12)	(0.64)	(0.03)	(0.01)	(1.38)
Age 20-24	0.07	-0.33	-0.07	-0.00	-1.69
	(0.05)	(0.46)	(0.05)	(0.01)	(1.39)
Age 25-29	0.08*	-0.35	-0.08	0.01	-0.98
	(0.04)	(0.51)	(0.07)	(0.01)	(1.75)
Age 30-34	0.12	0.12	-0.12	-0.02	-2.95
	(0.07)	(0.60)	(0.09)	(0.02)	(2.32)
Age 35-49	0.17**	0.79	-0.10	-0.00	0.69
	(0.07)	(0.61)	(0.07)	(0.01)	(2.74)
Female	-0.25***	0.03	-0.05**	-0.00	1.27
	(0.07)	(0.36)	(0.02)	(0.02)	(2.96)
Education: primary	0.01	0.03	-0.08	0.01	1.80*
to below SLC	(0.05)	(0.46)	(0.06)	(0.01)	(0.87)
Education: SLC and	-0.01	-0.16	-0.07	0.02	5.86**
beyond	(0.06)	(0.65)	(0.06)	(0.02)	(2.59)
Never married	-0.01	-0.36	0.01	-0.04	-3.00
	(0.04)	(0.34)	(0.06)	(0.03)	(2.04)
Brahmin and Chhetri	-0.06	0.60	0.03	0.00	-1.25
	(0.05)	(0.62)	(0.06)	(0.01)	(2.49)
Prior training	0.07	-0.30	0.03	0.00	-1.07
participation	(0.06)	(0.66)	(0.05)	(0.01)	(1.01)
Y_{t-1}	0.03	0.02	0.22*		0.07
	(0.04)	(0.04)	(0.11)		(0.06)
Constant	0.72***	5.23***	0.17**	0.02	7.37**
	(0.12)	(1.05)	(0.08)	(0.02)	(3.44)
Adjusted R^2	0.17	0.10	0.14	0.11	0.11
N	452	403	394	428	428
b. With regression adjustment					
Treatment	0.11***	0.59*	0.08***	-0.00	1.07
	(0.04)	(0.33)	(0.02)	(0.01)	(1.08)
N	452	403	394	428	428
c. With inverse probability weighting					
Treatment	0.11***	0.59*	0.08***	-0.00	1.07
	(0.04)	(0.33)	(0.02)	(0.01)	(1.08)
N	452	403	394	428	428
d. With selection of covariates using Lasso					
Treatment	0.10***	0.58*	0.08***	0.00	0.90
	(0.04)	(0.33)	(0.02)	(0.01)	(1.24)
N	452	403	394	428	428
e. Lee bounds					
lower	0.09*	0.10	-0.01	-0.01	-1.35
	(0.05)	(0.41)	(0.06)	(0.01)	(1.53)
upper	0.19***	0.99**	0.09***	0.00	1.24
	(0.06)	(0.43)	(0.03)	(0.01)	(1.43)
N	512	512	512	512	512
f. Significance level with randomization inference					
Treatment	0.18***	1.26***	0.05*	0.00	-2.09
N	452	403	394	428	428

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" indicating whether a trial participant was gainfully employed included home cultivation, a proxy for subsistence farming.

Table A.7: **ITT effect on the applicants left out by trainers**
(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	-0.15	0.01	0.02	-3.04
	(0.07)	(0.69)	(0.03)	(0.02)	(2.19)
Age 20-24	0.13	-0.35	0.01	0.02	3.30
	(0.09)	(0.98)	(0.07)	(0.02)	(2.77)
Age 25-29	0.23**	-0.23	-0.01	0.04	1.84
	(0.09)	(0.77)	(0.05)	(0.04)	(3.04)
Age 30-34	0.12	-0.56	-0.03	-0.00	0.73
	(0.12)	(1.13)	(0.07)	(0.03)	(2.86)
Age 35-49	0.25**	0.54	-0.01	0.01	2.60
	(0.10)	(1.26)	(0.06)	(0.02)	(2.74)
Female	-0.26***	-1.93**	-0.06*	0.00	0.74
	(0.04)	(0.70)	(0.03)	(0.04)	(2.11)
Education: primary	0.07	-0.50	0.01	0.01	-1.96
to below SLC	(0.05)	(0.59)	(0.02)	(0.04)	(2.09)
Education: SLC and	-0.03	-0.98	0.04	0.01	0.13
beyond	(0.06)	(0.88)	(0.05)	(0.01)	(2.01)
Never married	0.08	-0.06	0.00	-0.01	-0.29
	(0.09)	(0.55)	(0.03)	(0.03)	(2.75)
Brahmin and Chhetri	0.02	-0.01	-0.03	-0.03*	-0.20
	(0.09)	(0.81)	(0.02)	(0.02)	(2.01)
Prior training	0.04	-0.20	0.12	-0.02	-1.22
participation	(0.08)	(0.77)	(0.08)	(0.02)	(2.32)
Y_{t-1}	0.14**	0.11	0.12		0.21*
	(0.06)	(0.10)	(0.21)		(0.11)
Constant	0.55***	6.47***	0.04	-0.00	4.55*
	(0.11)	(1.05)	(0.06)	(0.03)	(2.35)
Adjusted R^2	0.24	0.10	-0.02	-0.04	0.11
N	240	209	203	229	229
b. With regression adjustment					
Treatment	-0.05	-0.63	0.01	0.02	-0.19
	(0.05)	(0.42)	(0.02)	(0.02)	(1.31)
N	240	209	203	229	229
c. With inverse probability weighting					
Treatment	-0.05	-0.63	0.01	0.02	-0.19
	(0.05)	(0.42)	(0.02)	(0.02)	(1.31)
N	240	209	203	229	229
d. With selection of covariates using Lasso					
Treatment	-0.05	-0.60	0.02	0.02	-0.76
	(0.05)	(0.41)	(0.03)	(0.02)	(1.29)
N	240	209	203	229	229
e. Lee bounds					
lower	-0.02	-0.88	0.02	0.02	-1.02
	(0.06)	(0.56)	(0.03)	(0.02)	(1.46)
upper	-0.00	-0.27	0.05**	0.03*	1.03
	(0.07)	(0.59)	(0.02)	(0.01)	(1.81)
N	290	290	290	290	290
f. Significance level with randomization inference					
Treatment	0.02	-0.15	0.01	0.02	-3.04**
N	240	209	203	229	229

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" indicating whether a trial participant was gainfully employed included home cultivation, a proxy for subsistence farming.

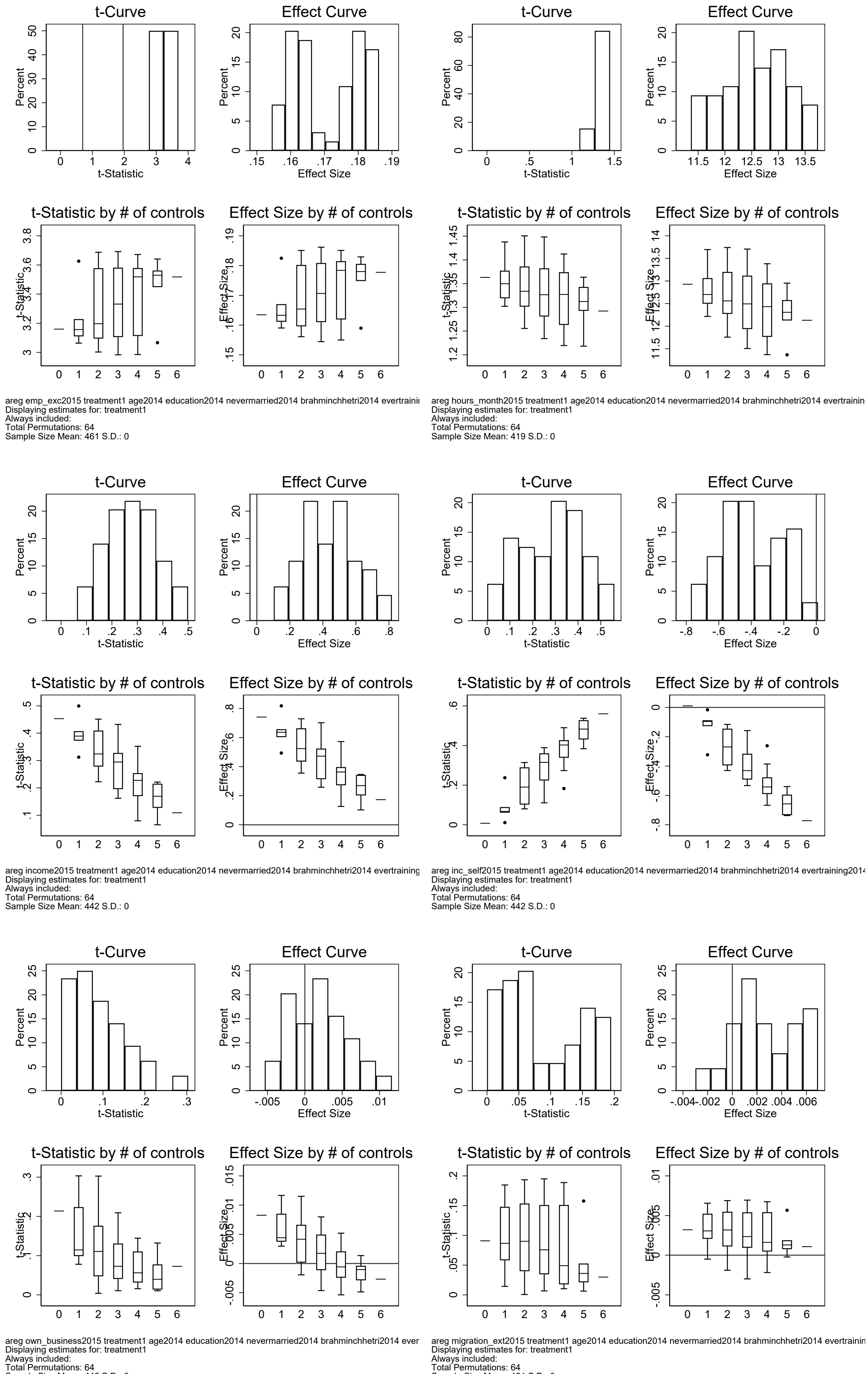

FIGURE A.1: **Specification check for p-hacking for the randomly selected group**

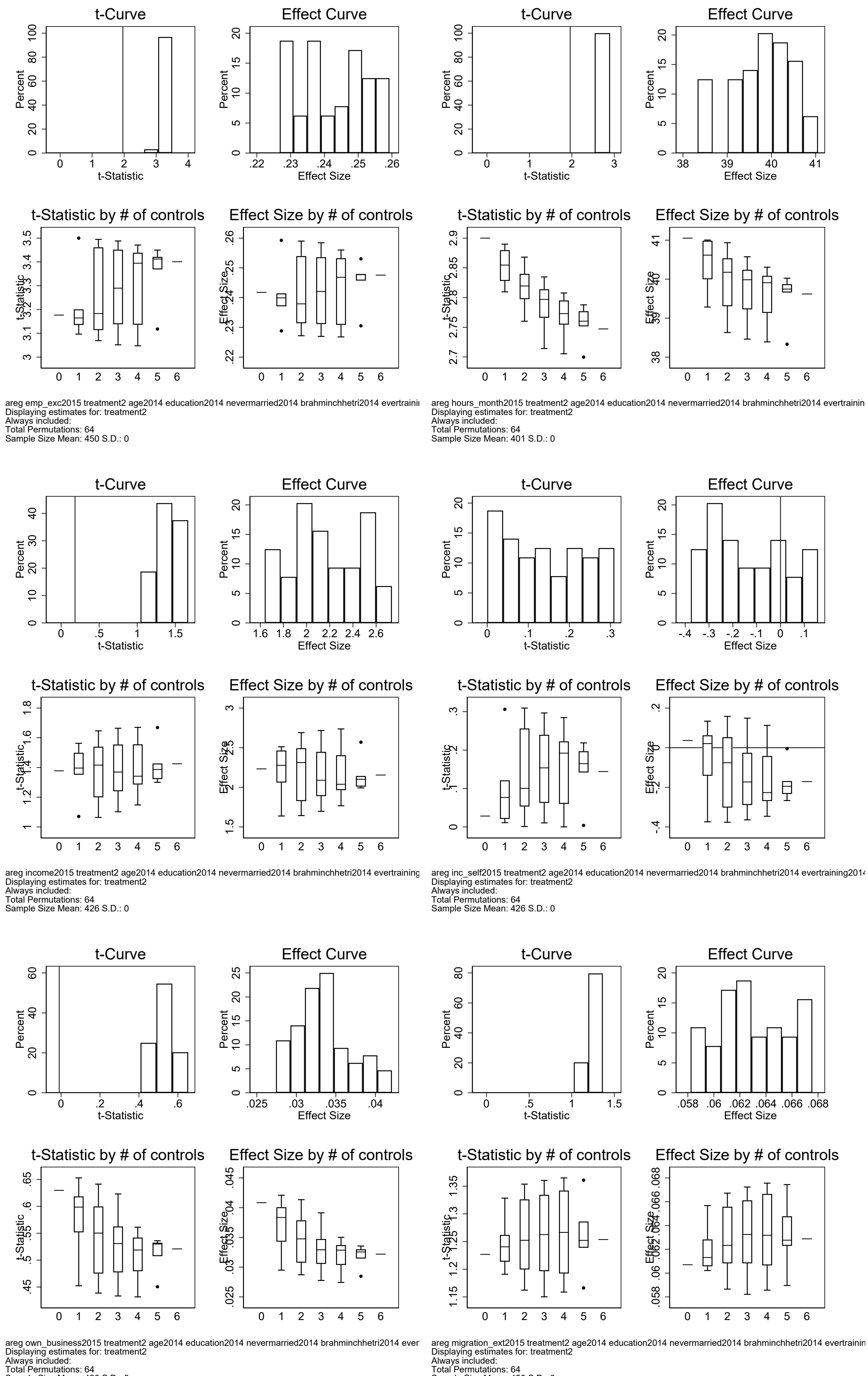

FIGURE A.2: **Specification check for p-hacking for the trainer selected trainee group**

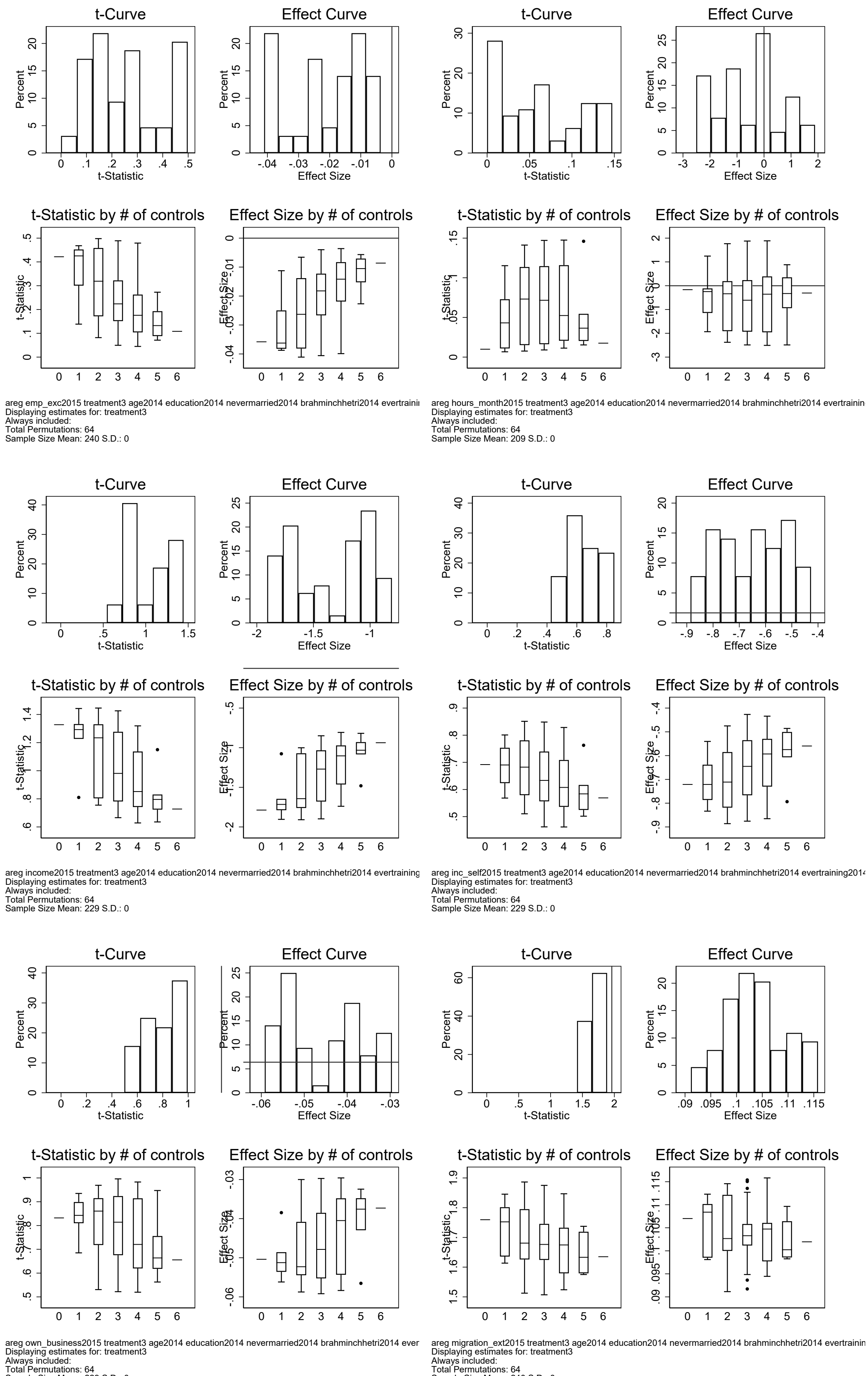

Figure A.3: **Specification check for p-hacking for the applicants left out by trainers**

Table A.8: **Minimum detectable effect size (MDES) by outcomes and group type**

Variable	Actual mean		MDES	Sample size	
	Control	Treatment	MDES	Control	Treatment
	(1)	(2)	(3)	(4)	(5)
a. Randomly selected trainees					
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 selected trainees					
Gainfully employed	0.61	0.80	0.14	114	338
Monthly hours worked	171.07	188.42	28.09	103	300
Monthly own income	8.46	9.77	5.06	112	316
Income working for oneself	3.07	3.39	3.96	112	316
Owns business	0.14	0.24	0.12	112	316
International Migration	0.15	0.15	0.11	114	338
c. Applicants left out by trainers					
Gainfully employed	0.61	0.55	0.18	114	126
Monthly hours worked	171.07	153.12	35.21	103	106
Monthly own income	8.46	4.87	4.53	112	117
Income working for oneself	3.07	1.49	3.72	112	117
Owns business	0.14	0.12	0.13	112	117
International Migration	0.15	0.20	0.14	114	126

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 group. 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.9: **LATE on randomly selected trainees**
(Corresponds to Table 3)

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Participated	0.23**	19.03	0.35	-0.76	-0.01	0.00
	(0.10)	(13.30)	(1.76)	(1.12)	(0.07)	(0.05)
Age 20-24	0.02	-12.04	1.86	0.01	0.04	-0.01
	(0.06)	(20.19)	(1.78)	(0.74)	(0.03)	(0.04)
Age 25-29	0.04	-12.49	1.50	1.98	0.13***	-0.09
	(0.10)	(24.39)	(1.89)	(1.29)	(0.04)	(0.07)
Age 30-34	-0.05	-1.61	6.63	7.50	0.14***	-0.08
	(0.10)	(27.16)	(5.02)	(4.88)	(0.05)	(0.07)
Age 35-49	-0.07	-21.09	-1.68	-0.14	0.14**	-0.10
	(0.14)	(25.47)	(2.51)	(2.11)	(0.07)	(0.08)
Female	-0.29***	-15.98	-10.70*	-5.43	-0.03	-0.15**
	(0.09)	(19.14)	(6.00)	(6.42)	(0.06)	(0.06)
Education: primary	0.08	-12.24*	3.46	4.43*	0.14***	-0.04
to below SLC	(0.05)	(7.29)	(2.53)	(2.43)	(0.05)	(0.05)
Education: SLC and	0.11**	5.33	5.30	6.29*	0.18**	0.00
beyond	(0.05)	(11.08)	(3.35)	(3.31)	(0.07)	(0.06)
Never married	-0.05	-26.82**	-1.96	-1.06	-0.07**	0.06
	(0.08)	(12.05)	(1.44)	(1.09)	(0.03)	(0.07)
Brahmin and Chhetri	-0.06	-9.72	3.98	5.37	0.05	-0.00
	(0.04)	(8.99)	(3.96)	(4.04)	(0.03)	(0.06)
Prior training	-0.11	-14.58	-3.73**	-1.89	0.05	-0.00
participation	(0.11)	(16.18)	(1.88)	(2.16)	(0.05)	(0.05)
Y_{t-1}	0.18***	0.09	0.04	-0.11*	0.07	0.13
	(0.05)	(0.06)	(0.05)	(0.06)	(0.09)	(0.16)
Constant	0.84***	197.18***	10.38**	1.10	-0.09	0.22**
	(0.14)	(27.98)	(4.37)	(4.36)	(0.08)	(0.11)
Adjusted R^2	0.17	0.12	0.11	0.16	0.23	0.24
N	461	419	442	442	442	461

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.

Table A.10: **LATE on trainer selected trainees**
(Corresponds to Table 4)

	Gainfully employed (1)	Monthly hours worked (2)	Monthly own income (3)	Income working for oneself (4)	Owns business (5)	International Migration (6)
Participated	0.39**	52.34**	3.89	-0.03	0.02	0.10*
	(0.17)	(23.90)	(3.38)	(1.63)	(0.11)	(0.06)
Age 20-24	-0.00	-9.58	1.74	-0.81	-0.03	0.04
	(0.05)	(13.31)	(1.31)	(0.93)	(0.05)	(0.04)
Age 25-29	-0.02	-12.26	-0.24	-0.92	-0.03	-0.02
	(0.05)	(14.84)	(1.99)	(1.44)	(0.06)	(0.04)
Age 30-34	0.02	-1.21	-0.85	0.77	0.07	-0.09*
	(0.06)	(17.75)	(1.88)	(1.57)	(0.06)	(0.05)
Age 35-49	0.06	23.90	1.46	2.29	0.13*	-0.02
	(0.07)	(18.38)	(2.55)	(2.40)	(0.08)	(0.03)
Female	-0.27***	-7.30	-8.36**	-2.27	0.13**	-0.17***
	(0.05)	(11.20)	(3.47)	(2.47)	(0.06)	(0.03)
Education: primary	-0.09*	-0.91	-1.45	-0.48	0.02	0.03
to below SLC	(0.05)	(12.69)	(1.48)	(1.26)	(0.04)	(0.03)
Education: SLC and	-0.06	-7.24	1.75	-0.07	-0.00	0.02
beyond	(0.07)	(19.39)	(2.86)	(2.30)	(0.05)	(0.05)
Never married	-0.09	-12.16	-4.97**	-2.35	-0.02	-0.05
	(0.07)	(8.95)	(2.29)	(2.35)	(0.05)	(0.04)
Brahmin and Chhetri	0.01	20.03	3.02*	1.08	0.07	-0.04
	(0.06)	(18.67)	(1.60)	(1.47)	(0.06)	(0.04)
Prior training	0.07	-9.01	-1.29	1.30	0.06	-0.08*
participation	(0.07)	(19.36)	(1.22)	(1.00)	(0.05)	(0.05)
Y_{t-1}	0.13**	0.03	0.08	0.04	0.21***	-0.09
	(0.06)	(0.04)	(0.07)	(0.05)	(0.06)	(0.12)
Constant	0.68***	172.78***	8.17***	2.93	-0.09	0.11**
	(0.15)	(25.16)	(2.06)	(2.15)	(0.13)	(0.05)
Adjusted R^2	0.10	0.06	0.11	0.14	0.28	0.07
N	452	403	428	428	428	452

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.

TABLE A.11: **LATE on applicants left out by trainers**
(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)
Participated	0.23	41.02	11.86	5.11	0.34	-0.76
	(0.63)	(166.26)	(12.77)	(7.09)	(0.21)	(0.51)
Age 20-24	0.09	-11.67	-0.01	-1.32**	0.01	-0.04
	(0.09)	(28.30)	(1.89)	(0.63)	(0.04)	(0.07)
Age 25-29	0.16*	-7.64	0.61	1.66	0.17***	-0.11
	(0.09)	(21.95)	(2.32)	(1.82)	(0.06)	(0.08)
Age 30-34	0.00	-22.09	-1.20	1.83	0.16**	-0.13
	(0.18)	(45.26)	(2.29)	(1.77)	(0.07)	(0.10)
Age 35-49	0.11	18.15	1.41	2.26	0.30**	-0.09
	(0.15)	(34.61)	(2.53)	(2.25)	(0.14)	(0.16)
Female	-0.33***	-60.97***	-8.55**	-3.35	-0.04	-0.16**
	(0.09)	(22.02)	(4.24)	(4.03)	(0.09)	(0.07)
Education: primary	-0.04	-12.50	1.72	1.32	0.08	-0.06
to below SLC	(0.08)	(20.75)	(2.34)	(2.05)	(0.07)	(0.06)
Education: SLC and	-0.08	-28.96	0.12	0.20	0.08	-0.01
beyond	(0.12)	(24.32)	(2.53)	(2.05)	(0.07)	(0.07)
Never married	-0.06	-2.07	-3.12	-1.11	0.01	0.00
	(0.09)	(15.20)	(1.93)	(0.98)	(0.06)	(0.05)
Brahmin and Chhetri	-0.10	-0.89	-0.59	0.83	-0.01	0.01
	(0.10)	(21.93)	(1.71)	(1.38)	(0.06)	(0.03)
Prior training	-0.03	-9.38	-0.38	2.33	0.05	-0.03
participation	(0.10)	(28.75)	(2.40)	(1.52)	(0.09)	(0.07)
Y_{t-1}	0.26***	0.13	0.25**	0.00	0.23***	0.24
	(0.07)	(0.13)	(0.12)	(0.06)	(0.07)	(0.21)
Constant	0.37***	170.97***	7.55*	1.24	-0.15	0.36***
	(0.13)	(32.13)	(4.24)	(3.93)	(0.10)	(0.10)
Adjusted R^2	0.21	0.10	0.08	0.08	0.14	0.03
N	240	209	229	229	229	240

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 randomly selected trainees 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)	0.14*	3.44	0.77	0.73	-0.02	-0.02
	(0.07)	(14.70)	(2.40)	(1.68)	(0.05)	(0.06)
Adjusted R^2	0.12	0.18	0.22	0.33	0.13	0.18
N	291	253	272	272	272	291
Treatment (female)	0.21	23.68	1.85	-0.46	0.02	0.01
	(0.17)	(19.66)	(1.52)	(0.90)	(0.10)	(0.02)
Adjusted R^2	0.17	0.15	0.11	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 trainer selected trainees 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	28.77	4.37	1.24	-0.10	0.06
	(0.12)	(17.44)	(4.10)	(2.53)	(0.12)	(0.07)
Adjusted R^2	0.09	0.09	0.11	0.23	0.19	0.04
N	267	226	246	246	246	267
Treatment (female)	0.29	58.56	1.86	0.39	0.08	0.08
	(0.35)	(42.68)	(1.87)	(1.24)	(0.05)	(0.05)
Adjusted R^2	0.17	0.14	0.12	0.15	0.34	-0.04
N	185	177	182	182	182	185

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 applicants left out by trainers 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)	-0.01	0.19	-0.65	0.14	-0.08	0.13
	(0.13)	(42.64)	(2.58)	(1.54)	(0.08)	(0.12)
Adjusted R^2	0.08	0.09	0.17	0.31	0.12	0.03
N	160	131	149	149	149	160
Treatment (female)	-0.03	-25.23	-1.32	-0.54	-0.03	0.00
	(0.08)	(23.49)	(1.61)	(1.45)	(0.10)	(.)
Adjusted R^2	0.18	0.19	0.28	0.36	0.39	.
N	80	78	80	80	80	80

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 randomly selected trainees 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	0.15	28.03	0.25	-2.91	-0.00	0.06*
(no education)	(0.15)	(21.12)	(3.80)	(3.74)	(0.05)	(0.03)
Adjusted R^2	0.13	0.11	0.10	0.09	0.07	0.65
N	129	123	124	124	124	129
Treatment	0.12	4.31	0.92	0.46	-0.11	-0.02
(primary education)	(0.11)	(14.99)	(3.72)	(2.68)	(0.06)	(0.07)
Adjusted R^2	0.22	0.15	0.14	0.16	0.43	0.15
N	203	184	193	193	193	203
Treatment	0.24	-0.68	-0.67	0.70	0.05	-0.04
(secondary education)	(0.20)	(21.76)	(3.88)	(1.93)	(0.09)	(0.18)
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.5: **ITT effect on trainer selected trainees 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	0.30	39.15	3.64**	0.78	0.00	0.13
(no education)	(0.31)	(37.98)	(1.70)	(0.95)	(0.07)	(0.10)
Adjusted R^2	0.16	0.13	0.14	0.17	0.34	-0.05
N	120	114	117	117	117	120
Treatment	0.14	35.95	0.58	-2.24**	-0.09	0.06
(primary education)	(0.15)	(27.10)	(2.03)	(0.93)	(0.13)	(0.07)
Adjusted R^2	0.14	0.08	0.11	0.20	0.33	0.04
N	220	188	201	201	201	220
Treatment	0.39*	17.02	8.22	2.86	0.20	0.18
(secondary education)	(0.19)	(17.69)	(8.38)	(4.53)	(0.11)	(0.12)
Adjusted R^2	0.30	0.22	-0.02	0.13	0.17	0.22
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.6: **ITT effect on applicants left out by trainers 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	0.00	-16.71	-2.69	-0.63	0.01	0.15
(no education)	(0.09)	(19.55)	(2.86)	(2.48)	(0.05)	(0.10)
Adjusted R^2	0.35	0.16	0.13	0.09	0.22	0.40
N	69	67	69	69	69	69
Treatment	-0.11	3.37	-0.65	-1.13	-0.17**	0.09
(primary education)	(0.11)	(38.53)	(2.81)	(1.57)	(0.07)	(0.10)
Adjusted R^2	0.16	0.09	-0.05	0.18	0.30	-0.01
N	108	90	98	98	98	108
Treatment	0.15	-84.91	0.10	-1.68	-0.10	0.27**
(secondary education)	(0.32)	(58.22)	(6.03)	(3.47)	(0.18)	(0.10)
Adjusted R^2	0.03	0.34	0.33	0.05	0.19	0.28
N	63	52	62	62	62	63

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 on randomly selected trainees 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	0.24*	20.02	1.45	0.87	0.02	-0.01
(low income)	(0.11)	(15.07)	(1.81)	(1.28)	(0.08)	(0.05)
Adjusted R^2	0.14	0.10	0.09	0.19	0.31	0.26
N	281	253	271	271	271	281
Treatment	0.07	7.14	-0.32	-0.79	-0.02	0.02
(high income)	(0.08)	(15.02)	(3.44)	(2.83)	(0.07)	(0.08)
Adjusted R^2	0.12	0.23	0.08	0.14	0.08	0.17
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.8: **ITT effect on trainer selected trainees 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*	49.47	2.30	0.07	0.08	0.05
(low income)	(0.19)	(29.93)	(1.72)	(0.81)	(0.09)	(0.06)
Adjusted R^2	0.19	0.05	0.20	0.13	0.21	0.15
N	296	261	281	281	281	296
Treatment	0.11	13.00	2.11	0.21	-0.14	0.07
(high income)	(0.13)	(18.29)	(3.60)	(4.03)	(0.10)	(0.07)
Adjusted R^2	0.09	0.21	0.00	0.11	0.44	0.01
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.9: **ITT effect on applicants left out by trainers 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	-0.07	-14.73	-2.04	-0.55	-0.08	0.05
(low income)	(0.11)	(25.48)	(2.03)	(1.14)	(0.06)	(0.06)
Adjusted R^2	0.13	0.13	0.16	-0.04	-0.01	0.19
N	149	124	141	141	141	149
Treatment	-0.09	4.10	-3.14	-1.97	-0.04	0.14
(high income)	(0.09)	(43.86)	(3.45)	(3.07)	(0.07)	(0.09)
Adjusted R^2	0.25	0.10	0.11	0.16	0.05	-0.07
N	91	85	88	88	88	91

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.10: **ITT effect of trainings 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 ×	0.07	24.04	1.10	-1.33	0.05	0.05
female	(0.17)	(23.60)	(2.69)	(2.02)	(0.12)	(0.06)
Treatment ×	0.01	-17.74	2.84	4.65	-0.08	-0.07
primary	(0.15)	(28.25)	(6.45)	(5.61)	(0.07)	(0.07)
Treatment ×	0.08	-19.67	2.28	5.94	0.11	-0.10
secondary	(0.21)	(29.61)	(6.12)	(4.97)	(0.09)	(0.14)
Treatment ×	0.02	7.31	1.70	1.03	-0.04	-0.01
high income	(0.06)	(9.62)	(1.30)	(1.22)	(0.05)	(0.04)
b. Trainer selected trainees						
Treatment ×	0.02	34.51	1.15	1.19	-0.00	0.05
female	(0.23)	(31.84)	(5.89)	(5.45)	(0.13)	(0.09)
Treatment ×	-0.14	11.04	-0.17	-0.73	-0.13	0.00
primary	(0.16)	(27.27)	(3.46)	(2.43)	(0.09)	(0.08)
Treatment ×	-0.15	-19.56	4.08	3.13	-0.02	-0.16*
secondary	(0.20)	(29.78)	(6.62)	(4.13)	(0.09)	(0.09)
Treatment ×	0.03	-6.32	2.51	0.95	-0.03	-0.03
high income	(0.06)	(12.54)	(1.71)	(1.27)	(0.05)	(0.06)
c. Applicants left out by trainers						
Treatment ×	-0.02	-30.10	7.76	3.99	-0.09	-0.03
female	(0.13)	(38.94)	(5.70)	(5.54)	(0.13)	(0.14)
Treatment ×	-0.09	12.06	2.56	0.89	-0.22	0.00
primary	(0.13)	(30.81)	(3.83)	(3.18)	(0.15)	(0.12)
Treatment ×	-0.28	-47.76	-3.01	0.41	-0.15	-0.20
secondary	(0.20)	(34.63)	(3.48)	(2.44)	(0.12)	(0.14)
Treatment ×	0.15	26.65	1.00	0.65	0.07	-0.01
high income	(0.11)	(24.89)	(2.48)	(1.82)	(0.05)	(0.06)

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34 and 35-49 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.11: **Estimated OLS coefficients and the corresponding ATE, ATT and ATU**

Outcome Variable	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	14.41	17.21	18.46	13.37
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	0.27	0.28	0.27	0.31
Monthly hours worked	37.80	37.80	37.83	37.70
Monthly own income	2.76	2.35	4.65	-4.14
Income working for oneself	-0.02	-0.37	1.72	-6.28
Owns business	0.01	-0.00	0.08	-0.22
International Migration	0.07	0.07	0.07	0.05
c. Applicants left out by trainers				
Gainfully employed	-0.03	-0.07	0.03	-0.18
Monthly hours worked	-4.59	-11.70	8.73	-32.72
Monthly own income	-1.41	-2.83	0.52	-6.33
Income working for oneself	-0.60	-1.08	0.09	-2.31
Owns business	-0.04	-0.06	-0.01	-0.11
International Migration	0.09	0.06	0.12	0.00

Note: The models also control for age (groups 15-19, 20-24, 25-29, 30-34 and 35-49 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 (Słoczyński, 2022).