



PRE-ANALYSIS PLAN

FOR

A Way Out? Apprenticeship Training, Employment and Social Transformation in Northern Nigeria¹

July 2019

Benjamin Crost
Dept. of Economics
University of Illinois

Oeindrila Dube
Harris School of Public
Policy, University of
Chicago

Marcus Holmlund
Research Department
The World Bank

Eric Mvukiyehe
Research Department
The World Bank

Abstract

Can improving labor market opportunities for marginalized youth alter their social and economic trajectories? We examine this question in the context of Northern Nigeria, a region with a long history of religious tension and violence along Christian-Muslim lines. Using a RCT, we study whether providing skills training to impoverished youth, who largely lack access to formal education and attend Islamic religious schools, can improve their engagement in income generating activities and reduce their participation in religious and political violence. We also examine whether training can promote gender empowerment, both among trained girls as well as their caretakers. The skills training is delivered through “Mafita”, a DfiD-funded initiative implemented by Adam Smith International (ASI). The training we examine in the impact evaluation described here takes the form of an apprenticeship, which provides on the job training in which youth train directly with skilled master craftsmen. The study involves 5,165 subjects and spans a two-year period, with endline data collected over the November 2018-May 2019 period. This document specifies the analysis plan for examining the effects of this training initiative, delineating the econometric specifications and outcomes we plan to examine, which, among others include: employment, income,

¹ Many thanks to Adam Smith International (Mafita Project Team) and DFID for making this work possible. Funding from DFID Nigeria is gratefully acknowledged. Also, thanks to the DIME team in DC and Nigeria for their invaluable inputs in making this study possible.

female empowerment, female confidence, participation in religious and political violence, religious extremism, social networks, and subjective wellbeing.

Keywords: *Skills development; job training; apprenticeship schemes; religious violence; political violence; field experiment; Northern Nigeria.*

1. Introduction

Fragile and conflict-affected states often have large reservoirs of unemployed or underemployed youth whose poverty and lack of economic opportunities puts them at risk of radicalization. Conflict and fragility can in turn lead to an increase in unemployment and poverty – a dynamic that can trap countries in a vicious cycle of poverty, unemployment and conflict, with devastating consequences for human development and well-being.

The UK Department for International Development (DFID) has funded the *Mafita* Program which delivers skills training interventions, implemented by Adam Smith International, that aim to improve the economic opportunities of marginalized youth in Northern Nigeria. Since the late 1970s, Northern Nigeria has been the site of religious and ethnic conflicts that have led to substantial violence and instability. These conflicts are thought to be fueled by the region's high rates of poverty and unemployment, as well as its large population of marginalized urban youth. Particularly notable in this context are the Almajirai, children of poor families from rural areas who migrate to cities to study in Qur'anic Schools. Observers of the region worry that, due to their lack of economic opportunities, these youths become easy targets for recruitment by violent groups and criminal organizations.

This study will evaluate the impact of an apprenticeship intervention delivered by the *Mafita* program to over 3,000 young men and women. The program is unique in targeting highly marginalized individuals, including many Almajirai. The intervention first delivered 6 months of foundational training in basic literacy, numeracy, project funding, working in teams, and managing relationships. It then linked participants to Master Craft Persons (MCPs) with whom they completed 6-9 months of apprenticeship.

The apprenticeship intervention recruited participants in Kano state. The program recruited from six categories of persons in the pilot period: Almajirai boys, Adolescent Girls in Islamic and Quranic Education (IQE), Persons with Disabilities (PWDs), Orphans and Vulnerable Children (OVC), and Early School Leavers (ESL). Recruitment was based on two primary selection criteria: (i) being in the age bracket 15 to 24 years old, and (ii) living on less than N225.76 (about \$0.65 in today's exchange rate) per day.

The proposed impact evaluation seeks to answer three questions.

1. *Can the apprenticeship training intervention delivered by the Mafita program increase the labor-market outcomes and economic well-being of marginalized youth?*
2. *Does this intervention induce changes in religious, social and political attitudes, and alter youth tendencies to engage in criminal and violent activities?*
3. *What—if any—heterogeneous impacts does the intervention produce?*

To answer these questions, the impact evaluation will make use of a randomized control trial that will randomly assign applicants into a treatment group that receives the intervention in the pilot phase and a control group that does not. The study will also investigate potential heterogeneous impacts of the intervention on a set of pre-identified dimensions.

1.1 Timeline of training intervention

From January 2017 to March 2018, the 3,066 participants were enrolled into the apprenticeship training program that took place in Kano, Nigeria. The apprenticeship training included 6 months of foundational training in basic literacy, numeracy, project funding, working in teams, and managing relationships, followed by 6-9 months of trade-specific on-the-job technical training that took place with clusters of Master Crafts Persons in particular sectors.

The trade-specific technical training was informed by the National Vocational Qualifications Frameworks standards, and participants had to complete assessments under this framework in order to graduate from the program.

After completion of the apprenticeship intervention, 900 participants (non-randomly selected by the Mafita program) were offered additional entrepreneurship training for two months.

1.2 Timeline of Survey

Prior to the start of the intervention, the Mafita program collected baseline data on a subset of program participants (as well as on other eligible youth who eventually were not enrolled in the study). This data was collected between October and December 2016 prior to the commencement of the intervention. Baseline data is available for approximately 80% of the study participants.

The World Bank collected endline data to assess program impact. This endline data collection commenced in December 2018 and was completed in May 2019. Data collection was preceded by several rounds of pilot testing and field testing of the survey instrument. The endline survey covers 5,165 participants in the apprenticeship study and 3,070 caregivers².

2. Randomization

Applicants were randomly assigned into a treatment group that received training starting in April 2017 and a control group that was excluded from training. The pool of applicants comprised 5,165 individuals,

² Unique caregivers were identified through the youth endline survey, where we collected detailed information about their caregivers, including caregiver demographics and contact information. This took the form of a caregiver listing exercise. The information collected from the youth survey was then used to generate the sampling frame for caregivers who were then interviewed.

3,066 of whom were randomly assigned into the treatment group while the remaining 2,099 were assigned to the control group.

The Mafita program asked to make sure that at least 40% of the treated individuals would be women, at least 20% would be Almajirai (Quranic school boys), and at least 20% would be Islamic Quranic Education (IQE) girls. To ensure this, as well as geographic balance across treatment and control, the intersection of geographical location, applicant type, and gender was used to define randomization strata. The geographical location variable used for this purpose was the local government area (LGA) of the applicant's residence. Applicant type was one of the following three categories; i) Almajirai boys, ii) IQE girls, and iii) other (PWD, OVC, ESL).

3. Outcomes

The study will estimate the effect of the apprenticeship intervention on a set of 7 primary and 9 secondary outcomes. The following list contains the outcomes of interest and all survey items grouped under each outcome.

Primary outcomes

1. Income generating activities (in past 30 days)
 - a. Worked in wage employment (t3>0)
 - b. Worked self-employed (u3>0)
 - c. Hours worked in wage employment (t4)
 - d. Hours worked in self-employment (u4)
 - e. Total income from wage employment (t12)
 - f. Total profits from self-employment (u9a)
2. Participation in political and religious violence. For our main analysis of this outcome, we will use responses from a self-completed audio module. We will also do a confirmatory analysis using responses to the same questions elicited in a list experiment (see Section 4.5 for details).
 - a. Used force or violence for a political cause (r13)
 - b. Participated in a group that sought to administer justice to someone for behaving badly/immorally (r12)
 - c. Used force or violence for reasons related to your religion (r9)
 - d. Participated in any riot, protest or demonstration that turned violent (r10)
3. Attitudes toward religious adherence and religious enforcement
 - a. Believes it's important that government ensures people follow religious laws (e6)
 - b. Vignettes on enforcing violations of religious rules. (For calculation of the index, the outcome will be defined as choosing either "report to authorities" or "beat

up” as the answer. In additional analysis, we will also report effects of treatment on the probability of choosing each individual answer.)

- i. Willing to punish alcohol consumption (e91)
- ii. Willing to punish miniskirt wearing (e101)
- iii. Willing to punish blasphemy (e111)

4. Female empowerment

- a. Desired/actual age of marriage / fertility (bc2, b2c, b2d, y32)
- b. Experienced domestic violence (o21-o215)
- c. In a marriage, who do you think should have the greater say in each of the following decisions (outcome defined as probability that decision is taken by wife or both partners equally) (y1-y10)
- d. Who decides how income will be used in household (y11)
- e. Would like daughters to work outside the home (y12)
- f. Various questions on women’s role in society (y14-y21)
- g. OK for a husband to beat his wife under various circumstances (y22-y28)
- h. OK for a wife to have her own opinion, even if it is different from her husband’s (y29)
- i. Would like to be working in 5 years’ time (ha1)
- j. Confidence in ability to solve Raven’s matrices (B11)

5. Caregiver female empowerment. In the primary hypothesis we will examine the sample of caregivers who have female wards. Treatment will be defined based on whether the caregiver has any female ward enrolled in Mafita. We will also conduct heterogeneity analysis with boys, to determine if the response of caregivers is significantly different for caregivers who have male wards enrolled in Mafita (as compared to those who have female wards enrolled in the program).

- a. Ok for unmarried and married women of ages 14, 18 and 22 to be earning an income inside home and outside home (b2a-e)
- b. Age at which women should get married (b3_1); Age at which women should have their first child (b3_3)
- c. Various questions on what women should be allowed to do in society and what boys and girls should have to do and have prioritized toward them (b5_1 – b5_8)
- d. Various questions on equality in household decision making (b6_1-b6_10)

6. Anti-social behavior. For our main analysis of this outcome, we will use responses from a self-completed audio module. We will also do a confirmatory analysis using responses to the same questions elicited in a list experiment (see Section 4.5 for details).

- a. Used tramadol in past 3 months (r2)
- b. Used kayan maye in past 3 months (r1)
- c. Stolen objects or money from someone, when that person was not present (r5)

- d. Stolen objects or money from someone, when that person was present (r6)
- e. Gotten into a fight where I tried to physically hurt someone (r3)
- f. Did some work for a criminal group (d'an ta'adda) in the past 6 months (r14)

7. Generosity toward other religions: donation experiment. See section 4.4 for details.

Secondary outcomes

- 8. Attitudes towards political violence
 - a. Use of violence is justified in protesting an injustice (I3)
 - b. It is good to use violence to resolve problems (I5)
 - c. It is ok to use force or violence for political cause (I6)
- 9. Religious social network (for the purpose of this outcome, a respondent's social network is defined as the five people he/she spends the most time with)
 - a. Number of people from other religion in social network (e22)
 - b. Trusts people from other religions (e45)
 - c. Trusts leaders from other religions (e57)
- 10. Religiosity
 - a. Time spent on religious activities (c22c)
 - b. Importance of religion in respondent's life (e35)
 - c. Other things are more important than religion (r16)
 - d. Often has a strong sense of God's presence (h211)
- 11. Subjective well-being
 - a. Cantril's Ladder (h11)
 - b. MHI-5 (mhi1-mhi5)
- 12. Self-esteem
 - a. 7 item index (h212, h31-h39)
- 13. Skills
 - a. Literacy assessment (la1-la5)
 - b. Numeracy assessment (na1-na7)
- 14. Assets and consumption
 - a. 12-item module (Q1-Q7, Q11-Q12)
 - b. Money spent on various categories (d22a-d22c, d22f-d22k)
- 15. Job search behavior
 - a. In the past 6 months, did you actively look for a job? (v1)

- b. How many months out of the past 6 months did you actively look for a job? (v2)
- c. In the past 30 days, did you actively look for a job? (v3)
- d. How many days out of the past 30 days did you actively look for a job? (v4)
- e. In the past 7 days, did you actively look for a job? (v5)
- f. How many days of the past 7 days did you spend actively looking for a job? (v6)
- g. In the past 6 months, did you try to start your own business (v7)?
- h. In the past 30 days, did you try to start your own business (v8)?

16. Range of social networks (for the purpose of this outcome, a respondent's social network is defined as the five people he/she spends the most time with)

- a. Number of people in social network who are currently employed (e24)
- b. Number of people in social network who live in same neighborhood as respondent (e26)

For the directly measured outcomes (all outcomes except 7) we will combine all survey responses related to an outcome into a mean effects index, following Kling et al. (2007). To do this, we will first express responses in terms of standard deviations from the control group mean. We will then sum all standardized responses related to an outcome into an index switching signs if necessary to ensure that the positive direction always indicates a “better” outcome. We will also report a robustness test using the method of Anderson (2008), which weights the index items by their inverse covariance matrix. We will also present estimates of individual indicators within each family to better gauge how various indicators contribute to overall effects within families.

For certain outcomes, we will test the primary hypothesis only for the subsample of male or female respondents. Specifically, we will test the primary hypothesis for outcomes 2 and 6 (participation in political/religious violence, and anti-social behavior) only for the subsample of male participants, since qualitative evidence suggests that the program's effect on these outcomes is likely to be substantially larger for male participants. In addition, we will test the primary hypothesis for outcome 4 for the subsample of female respondents since the outcome is less relevant for men in this context.

As part of the heterogeneity analysis, we will also report estimates for outcomes 2 and 6 for women and outcome 4 for men. (For the latter, we will focus on items 4.c-4.i since these measure attitudes towards gender equality and the role of women in society. We will not include items 4.a (desired age of fertility/marriage) and 4.b (experience with domestic violence), since these are less relevant for boys in this context.)

For outcome 7, which is measured by a choice experiment, we will use the estimation procedure described in section 4.4.

For each primary outcome, we will estimate one regression using the mean effects index as the dependent variable and adjust p-values for multiple comparisons by controlling the False Discovery Rate (FDR) using the step-up procedure proposed by Benjamini, Krieger and Yekutieli (2006). When examining individual indicators within outcomes we will control the FDR across indicators within the outcome.

4. Estimation

4.1 *Balance and compliance*

We will test balance on observables using several sources of data. First, we will use the observable characteristics collected during the Mafita recruitment process. These include applicants' age, gender, location of residence and primary and secondary trade requested. Since gender and location were used during the stratification, we will only test balance on age and trade choices, controlling for strata fixed effects.

Second, we will use data from a pre-treatment survey collected by the Mafita program. This survey was not part of the impact evaluation; its goal was merely to inform Mafita about the characteristics and motivations of applicants to the program. As part of this effort, the survey collected data on some outcomes similar to the ones we will study, as described below. One limitation of this data source is that not all of the applicants that formed the randomization sample were surveyed. We will therefore only test balance for the subset of approximately 75% of individuals in our sample, for whom data is available from this source.

Finally, we will test balance using data from our follow-up survey. For this, we will select a small set of variables that could not have been influenced by treatment, such as an individual's religion and ethnicity, and other plausibly fixed characteristics.

Finally, we will estimate the rate of compliance with the randomization by estimating the effect of treatment assignment on participation and completion of the Mafita program, as reported in the program's administrative data. We will do this using the same estimating equation as for our other outcomes of interest, described in Section 4.3.

4.2 *Attrition, outliers and missing data*

We will test the hypothesis that treatment does not affect the probability of attrition using a linear probability model that regresses an attrition indicator on the treatment indicator and a set of strata fixed effects. We will also report Lee bounds for all estimates, to account for possible effects of unbalanced attrition.

We will deal with outliers by winsorizing unbounded variables at the 99th percentile.

To deal with missing data on outcome variables, we will follow Kling et al. (2007) and impute missing values by setting them equal to the mean of the variable for the relevant treatment group. We will deal with missing data on control variables by setting the missing value to an arbitrary number (e.g. zero) and including a missing value indicator for each control variable that has missing values.

4.3 Estimation for directly measured outcomes

We will separately estimate the intention-to-treat (ITT) effects of the intervention on the outcomes listed above, using the following linear regression:

$$Y_i = \beta_0 + \beta_1 T_i + \gamma X_i + \varepsilon_i$$

where Y_i is the outcome of interest for individual i , T_i is an indicator for assignment to the treatment group and X_i is a vector of pre-program covariates. This vector will include the following variables: age, ethnicity, religion, and indicators for randomization strata (LGA-by-applicant-type). For outcomes for which baseline data is available from the baseline survey conducted by the Mafita program, we will also control for the baseline outcome using an ANCOVA specification.

As stated above, we will adjust for multiple comparisons by controlling the False Discovery Rate using the step-up procedure proposed by Benjamini, Krieger and Yekutieli (2006). In addition to reporting point estimates and confidence intervals of the coefficient β_1 , we will also report Lee bounds for the effects of attrition.

As mentioned above, our primary specification will estimate the intention-to-treat (ITT) effect. However, we will also report the Local Average Treatment Effect (LATE) from a regression that uses treatment assignment to instrument for participation in training (as measured by administrative data).

4.4 Donation experiment

Outcome 7 (generosity towards members of other religions) is measured by a choice experiment in which respondents can donate part of the compensation for taking the survey to one of three NGOs. For this experiment, respondents are randomly allocated to three groups. The first group is asked how much they would like to donate to Islamic Relief, an organization that works with disadvantaged Muslim youth. Respondents in the second and third group are asked how much they would like to donate to UNICEF or the Nigerian Red Cross, respectively, organizations that work with disadvantaged youth from all religions. To test whether the program affected participants' generosity towards other religions, we estimate the following regression on the subsample of Muslim respondents

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 T_i \times UNICEF_i + \beta_3 T_i \times RC_i + \delta_1 UNICEF_i + \delta_2 RC_i + \gamma X_i + \varepsilon_i$$

$UNICEF_i$ and RC_i are indicators for whether the respondent was randomized into the group asked to donate to UNICEF and the Red Cross, respectively. All other variables are defined as above. The coefficients of interest will be β_2 and β_3 , which capture whether treatment made respondents more likely to donate to organizations not associated with their own religion, relative to an organization associated with their own religion. The main hypothesis test for this outcome will be an F-test of joint significance of β_2 and β_3 .

Also, of interest, though not part of a pre-specified hypothesis, will be the coefficient β_1 , which captures the effect of treatment on donations to Islamic Relief. A positive coefficient would reflect that treatment increased the generosity of respondents toward disadvantaged youth of their own religion.

Since we expect that almost all respondents will be Muslim, we will not conduct a separate analysis for non-Muslim respondents, due to lack of statistical power.

4.5 List experiment

Some outcomes of interest of this study measure sensitive issues such as participation in religious and political violence (outcome 2) and anti-social behavior (outcome 6) and may therefore be subject to social desirability bias. To mitigate this bias, we elicited the questions related to these outcomes with a self-completed audio-module. In this module, respondents are played audio recordings of the questions on headphones and enter their responses on a tablet, out of view of the enumerator. However, there might be some remaining social desirability bias since respondents know that researchers who analyze the survey data will be able to see their responses.

To address this concern, we elicited answers to the same sensitive questions using a list experiment. In this survey experiment, respondents are read a list of statements and asked to reveal how many they agree with, but not which specific ones. Furthermore, respondents are randomly divided into two groups, one that is read a list of four non-sensitive statements and one that is read a list of the same four non-sensitive statements, plus the sensitive one we are interested in. The percentage of respondents that agree with the sensitive statements can be statistically estimated by taking the difference in the average number of statements the two groups agree with. However, it is impossible for the enumerator or researcher to know whether any specific respondent agreed with the sensitive statement, reducing concerns over social desirability bias.

We will use this method to estimate the program's effect on the items associated with outcomes 2 and 6, by estimating the following regression:

$$Y_{ik} = \beta_0 + \beta_1 T_i + \beta_2 Sens_{ik} + \beta_3 T_i \times Sens_{ik} + \gamma_1 X_i + \gamma_2 X_i \times Sens_{ik} + \varepsilon_{ik}$$

Y_{ik} is the number of statements agreed to by individual i for list k . $Sens_{ik}$ is an indicator for whether individual i was assigned to the group that was read the sensitive item in list k . All other variables are defined as above.

We will estimate this regression jointly for all items listed under an outcome. Thus, for outcome 2, the regression will contain four data points per individual, one for each of the four items listed under the outcome (2.a-2.d). Standard errors will be clustered at the individual level to account for possible within-individual correlation in answers. The coefficient β_3 will estimate the program's effect on the average probability of agreeing with the four sensitive items, and we will conduct a hypothesis test using this estimate. We will also report descriptive results of similar regressions for each of the four items separately.

To test for order effects, we randomized the order in which the audio module and list experiment were administered to survey participants. We will use this randomization to test whether responses to the list

experiment are affected by having previously answered the same questions in the audio module, and vice versa.

4.6 Effect of being assigned to preferred trade

As part of the enrollment process in the apprenticeship program, individuals were allowed to express preferences for specific trades they would like to receive training in. In particular, individuals were asked to list two trades that were their first and second priority. Unfortunately, it was not possible to assign every applicant to one of their preferred trades, since the number of positions in each trade was limited, and some trades were oversubscribed.

After randomizing individuals into treatment and control groups, we conducted an additional randomization to generate variation in the likelihood of being assigned to one's preferred trade. To do this, we randomly assigned ranks to all individuals in the treatment group and fulfilled their trade preference in order of ascending rank. Thus, the individual with rank 1 would definitely be assigned to their first priority trade and take up one slot in the trade.

After this, the individual with the next highest rank individual would be assigned to their first priority trade if there were still slots available. If not, they would be assigned to their second priority trade if there were still slots available in that trade. This algorithm would be repeated for every individual in order of their rank. At the end of the procedure, individuals who received neither their first nor second priority trade were randomly assigned a still available trade.

Our analysis will exploit the fact that individuals with high rank in this procedure were more likely to be assigned a preferred trade in an instrumental variables framework. We will estimate the following first stage regression using data from the treatment group only (since the control group was not assigned ranks or trades):

$$primary_i = \alpha_0 + \alpha_1 rank_i + \gamma X_i + \delta Z_i + \varepsilon_i$$

where $primary_i$ is an indicator for whether individual i was assigned to their primary preferred trade and $rank_i$ is the rank assigned to individual i in the allocation of trades. In addition to the usual control variables (X_i) listed above, we will also control for Z_i , a vector of interactions between indicators for the individual's location (LGA) and indicators for the individual's preferred primary trade. This adjusts for the fact that the number of available slots differed across trades and locations, so that individuals had a higher probability of being assigned their preferred trade if they chose a trade that was less popular or had more available slots. The coefficient α_1 captures the effect of being ranked one rank lower on the probability of receiving one's primary preferred trade.

We will then estimate the following reduced form regression:

$$Y_i = \beta_0 + \beta_1 rank_i + \gamma X_i + \delta Z_i + \varepsilon_i$$

The coefficient β_1 captures the effect of being ranked one rank lower on the outcome of interest.

Finally, we will estimate a two-stage least squares regression to estimate the effect of being assigned one's preferred trade:

$$Y_i = \delta_0 + \delta_1 \widehat{primary}_i + \gamma X_i + \delta Z_i + \varepsilon_i$$

where $\widehat{primary}_i$ is the predicted value of $primary_i$ from the first-stage regression. The coefficient δ_1 captures the effect of being assigned one's primary preferred trade on the outcome of interest. We will also estimate the same regression using as the endogenous variable an indicator for whether individual i was assigned either their primary or secondary preferred trade (this time also controlling for interactions between location and indicators for secondary trade choices).

In addition to the outcomes listed in Section 3 above, we will also estimate this regression on an indicator for whether the participant completed the training, to test whether being assigned to a preferred trade affects the probability of dropping out of the program.

We will also use the entire sample to estimate an instrumental variables regression where we instrument training in a preferred trade with the rank variable ($rank_i$) while controlling for the treatment indicator:

$$Y_i = \beta_0 + \beta_1 primary_i + \beta_2 T_i + \gamma X_i + \delta Z_i + \varepsilon_i$$

For this analysis, we will assign the control group a rank of zero, to deal with their missing values for this variable.

4.7 Heterogeneity analyses

We will test whether treatment effects are heterogeneous along the following three dimensions:

a. Gender (Male/Female)

As part of this analysis, we will estimate the effect of treatment on outcomes 2 and 6 (participation in political/religious violence and anti-social behavior) for girls and outcome 4 (female empowerment) for boys. As discussed above, we will focus on items 4.c-4.i when examining the latter effects. We will also examine heterogeneous effects on male and female caregivers for outcome 5 (caregiver female empowerment).

b. Youth category (almajirai versus other boys, IQE girls versus other girls)

There is a widespread perception in the country that outcomes 2 (participation in religious and political violence) and 6 (anti-social behavior) are higher among almajirai than among other males. If baseline levels of these outcomes are in fact higher for almajirai, we might expect the program to have a larger potential to reduce them. We will test this hypothesis by comparing the effect of the program on almajirai relative to other boys, with a particular focus on outcomes 2 and 6.

We will also test whether treatment effects are different for IQE girls and other girls.

For almajirai and IQE girls, we will also report effects on an extended version of outcome 9 that includes survey item e25 (number of people studying in religious school in social network) to test whether the intervention broadens their social networks beyond the religious schools they are affiliated with.

c. Role of entrepreneurship training

We will report estimates for all of our primary outcomes for a subsample that excludes individuals who received entrepreneurship training. Since entrepreneurship training was non-randomly assigned, this estimate should not be interpreted as the causal effect of the “basic” training without entrepreneurship training. However, this analysis is meant to provide some suggestive evidence for whether our results are largely driven by individuals who received entrepreneurship training.

d. Machine learning for heterogeneous effects

We will also attempt to detect heterogeneous effects across other dimensions by implementing the causal forests procedure described by Wager and Athey (2018).

References

- Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, 103(484), 1481-1495.
- Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up false discovery rate controlling procedures. *Biometrika*, 93(3), 491-507.
- Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.
- Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228-1242.