Entrepreneurship Education and Teacher Training in Rwanda
PIs: Moussa Blimpo, Todd Pugatch
Tracer Survey 2019 Analysis Plan
June 2, 2020

1 Overview

This document presents an analysis plan for the Educate! Exchange Program 2019 tracer survey of program alumni. The program ran in 103 randomly assigned Rwandan secondary schools from 2016-2018. Data were also collected from a randomly assigned control group of 104 schools. The 2019 tracer survey was in the field June and July 2019. The survey targeted all study participants from the 2016 baseline survey, which was a random sample of 15 students per school. All surveys were conducted by the Kigali office of Innovations for Poverty Action (IPA).

The 2019 tracer survey occurred in two interview formats: phone and in-person. First, enumerators attempted to contact respondents by phone, using phone numbers provided in previous rounds. Enumerators interviewed all willing participants reached by phone. A subsample of those interviewed by phone were selected for an additional in-person interview. The in-person interview covered the same questions as the phone interview, plus additional questions.

Due to survey budget constraints, initial screening for in-person interviews was based on the respondent's availability and proximity to Kigali. Enumerators then attempted to balance respondent gender and treatment arm within this eligible group. In-person interviews occurred at respondents' homes or other mutually agreed location. We discuss our approach to analyzing both survey formats in greater detail below.

2 Methodology

2.1 Main effects

Our analysis will build on the analysis plan for the 2018 endline survey, which was published as a Stage 1 Registered Report in *Journal of Development Economics* (Blimpo and Pugatch 2019). The main statistical method we use is ordinary least squares linear regression (OLS). This is the appropriate method because randomized control trials solve (in principle) the selection problem for estimation of the mean outcome difference due to assignment to treatment. Moreover, OLS allows us to adjust easily for the stratification of treatment, ensuring that we rely on experimental variation. We will cluster standard errors by school to account for correlated outcomes among students within a school, the unit of treatment.

The main results will come from the regression:

$$y_{isg} = \alpha + \beta T_{sg} + \delta y_{0isg} + \gamma_g + \varepsilon_{isg} \qquad (1)$$

where i indexes students; s indexes schools; and g indexes strata. The strata are district-school type cells, where school types are public and non-public. In equation (1), y is an outcome (with y_0 the outcome at baseline, if available); T is an indicator for assignment to treatment; γ is a stratum

fixed effect; and ε is an error term. Because randomization occurred within strata, the strata fixed effects ensure that treatment assignment T is unrelated to the error term. The coefficient of interest is β, which measures the intent to treat (ITT), or the effect of the offer of teacher training T on the mean outcome.

Table 1 lists the outcomes we plan to test using equation (1).

We will not impute missing values for any dependent variables. For covariates (e.g., baseline outcomes), we will replace missing values with the control group mean and include a dummy for missing values in the regression (Haushofer and Shapiro 2015). To deal with outliers, we will winsorize all financial variables (e.g., income, savings) at the 99th percentile. We will not impute values for outliers in other variables.

The estimation sample is all students sampled at baseline. Our main analysis will use responses to the phone survey, except when the outcome appears only in the in-person survey (see Table 1). All in-person respondents also completed a phone survey. We will therefore check robustness of results to using outcomes from the in-person survey. In these robustness checks, we will also include a dummy for interview format to account for systematic differences between them.

2.2 Heterogeneous treatment effects

The effect of the program may differ among students or teachers. Testing for such effects is important as it may point to key policy nuances. We will allow treatment effects to vary according to observable characteristics of a student or school by modifying equation (1) as:

$$y_{isg} = \alpha + \beta_1 T_{sg} + \beta_2 (T_{sg} \times X_{0isg}) + \beta_3 X_{0isg} + \delta y_{0isg} + \gamma_g + \varepsilon_{isg} \qquad (2)$$

where X_0 is some characteristic determined prior to the treatment. A non-zero value of β_2 indicates that the effect of treatment differs according to X_0.

To keep estimation tractable, we will limit estimation of equation (2) to the following student outcomes:

- university enrollment (U1˙univenroll)
- any business involvement [earn˙receive˙401e]
- employment [earn˙receive˙401e]
- total income from business and employment [howmuch˙401ae]

The characteristics X_0 we plan to test are the following:

- gender (BSQ301)
- past academic performance (S3 exam score, BSQ404)
- household socioeconomic status (SES): SES will be measured as an indicator for being above the median of the first principal component of the following variables:
    - parents' education (BSQ310)

- household assets (BSQ303-308)
- parents' occupation (BSQ309, indicator for business/professional)

3 References

Blimpo, Moussa P., and Todd Pugatch. 2019. "Entrepreneurship Education and Teacher Training in Rwanda." *Journal of Development Economics* 140 (September): 186–202. https://doi.org/10.1016/j.jdeveco.2019.05.006.

Haushofer, Johannes, and Jeremy Shapiro. 2015. "Household Response to Income Changes: Pre-Analysis Plan." https://www.socialscienceregistry.org/docs/analysisplan/384/document.

Table 1: Outcomes

Outcome	Variable	Baseline outcome	Notes
Attrition			
overall			
phone survey			
in-person survey			
Calendar activity			
secondary enrollment	ac0_educsecond	N/A	Separate outcomes by month, January-July 2019
university enrollment	ac1_educuniv	N/A	Separate outcomes by month, January-July 2019
training enrollment	ac2_eductrain	N/A	Separate outcomes by month, January-July 2019
business/management	ac3_bus_mgt	N/A	Separate outcomes by month, January-July 2019
employment (paid)	ac4_paid_employ	N/A	Separate outcomes by month, January-July 2019
employment (unpaid)	ac5_unpaid_employ	N/A	Separate outcomes by month, January-July 2019
Academic performance			
complete secondary	S1_complt_status	N/A	
S6 national exam score	S7_natexam_score	BSQ404	
university enrollment	U1_univenroll	N/A	
TVET enrollment	U1_univenroll	N/A	
Entrepreneurship & workplace skills			
entrepreneurial spirit	e1-e8	N/A	Based on Bruhn et al, *Journal of Political Economy* 2018
work skills index	ss1-ss8	N/A	From World Bank STEP Skills survey
creativity	use_pole	N/A	number of uses of a pole
monthly discount rate<100%	moneyoffer_801	BSQ801	in-person survey only
monthly discount rate<300%	moneyoffer_802	BSQ802	in-person survey only
started community project	comm_project_1101e	N/A	
Non-cognitive skills			
aspiration: university or beyond	schooling_1100	BSQ1100	
aspiration: business or professional	occup_1101	BSQ1101	
aspiration: create business	busnsftr_1102	BSQ1102	

belief in future	bf1-bf7	BSQ1200/1202/1204/1206/1208	Baseline outcome is locus of control
grit	g1-g12	BSQ1300-1303	Outcome is 12-item scale, while baseline outcome is 4-item scale.
Economic activity			
business participation: all	earn_receive_401e	BSQ600	
business participation: own business	type_busi_402	BSQ604	
business participation: student business club	type_busi_402	BSQ604	
business participation: family or peers	type_busi_402	BSQ604	
non-agricultural business	type_busi_407	BSQ601	
business has paid employees	paid_worker_406	BSQ600	
employed	earn_receive_401e	BSQ500	
income (total)	howmuch_401ae	BSQ503	
business profit	often_earn_409, lasttime_profits_410	BSQ608	
savings (any)	savedmon_804e	BSQ804	
savings (amount, conditional on any)	savedamt_805e	BSQ805	
borrowed money for business	borrow_800		in-person survey only