The effect of entrepreneurial training on business creation and profitability among the Ugandan youth

Whitney Mgbara, Shruti Bathia, David Contreras, William Krinsman

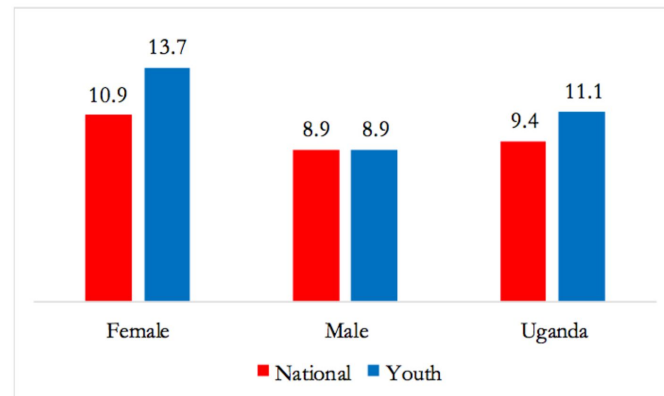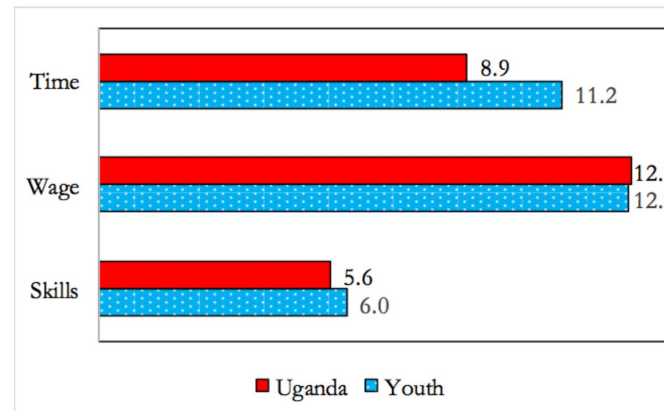# Background

Figure 1: Unemployment rate in Uganda-2012/13 (%)

Figure 2: Nature of underemployment in Uganda, 2012/13 (%)

Background

- Possible solution: Self-employment
- Governments recognize that their economy would benefit from better-trained entrepreneurs.
- What is needed?
  - Youth need the right combination of human, financial, and social capital to improve their welfare.
- Uganda and 22 other African countries have mainstreamed entrepreneurship training in high school through support from the ILO.

Current Literature
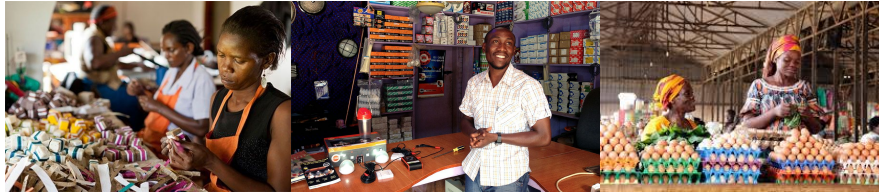
- "Business training leads to improvements in knowledge of good business practices"
  - Karlan and Valdivia (2011) and Valdivia (2011) in Peru, Drexler et al. (2011) in the Dominican Republic, Berge et al. (2011) in Tanzania.
- In Sri Lanka de Mel, McKenzie, and Woodruff (2012) examine the effects of an ILO business training program on business success of both existing female entrepreneurs and the general population of women.
- **Our project aims to expand this research to all classes of pre-entrepreneurs.**

Research Question

1. What are the effects of entrepreneurial training on youth participation in economic activities?
2. Does training influence the likelihood of starting a business after high school graduation?
3. Does training lead to an increase in total earnings?

_____ Photo taken during program training

The intervention: Entrepreneurial Training Program

Photo taken during program training

Location: Uganda

Recent high school graduates age 17–19

High quality 3 week in-residence mini MBA

Two versions
◦ 75% hard skills and 25% soft skills
◦ 75% soft skills and 75% hard skills

Both get business plan preparation

Large nationally representative sample

Follow-up 3+ years later

Study Attributes

- Target Population: youth in Uganda, post–high school graduation
- Sample: Students enrolled in the last year of high school in 4 regions of Uganda in 2013.
  - 40% in the West, 20% in Jinja, 20% in Mbale, and 20% in the North.
- A random sample of 4,400 students out of the eligible 7,421.
- Students and teachers assigned to one of three groups (hard skills, soft skills, control) randomly.
- On each step of the sampling process we stratified by school and gender.
- In the follow-up survey we were able to interview 3,891 participants (88.4%).

The Causal Model

- Endogenous nodes:
    - X= (W, A, Y)
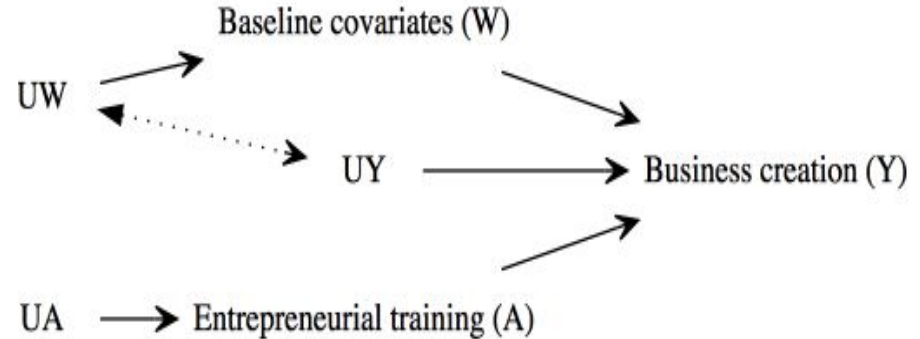
Note: W is a set of covariates and Y is a set of outcomes

- Exogenous nodes:
    - U= (UW, UA, UY)~PU.
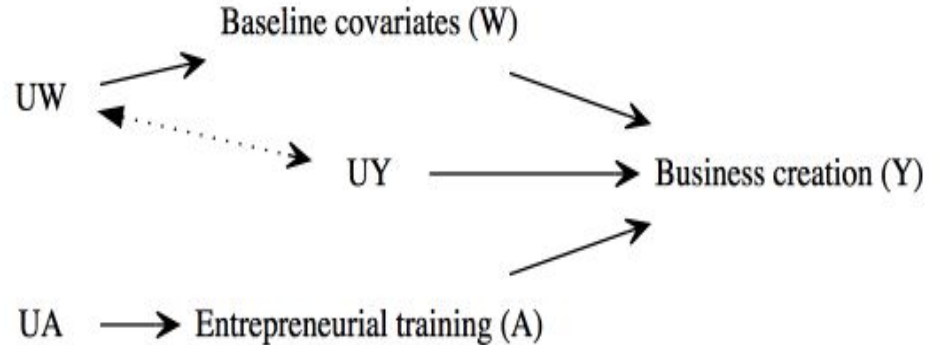- Structural equations F:
    - W =fW(UW)
    - A=fA(UA)
    - Y=fY(W, A, UY)

The Causal Model

- Intervention node
  - A
- Exclusion Restrictions
  - W to A
- Independence Assumption
  - UW to UA (vice versa)
- Positivity Assumptions

Variable type	Variable name	Description
Y	Y1 = Business creation Y2 = log(total earnings)	Y1 =1 if respondent started a business after graduation from high school Y2 = discounted sum of total earnings from profits and wage work
A	Treatment	A = 1 if individual participated in entrepreneurial training (hard and soft skills pooled)
W	W1 = Sociodemographic characteristics W2 = Cognitive skills W3 = Risk and time preferences W4 = Personality characteristics	Gender, age, parent's income source and education level, boarding student, perceived socioeconomic level, Raven score, math score, GPA, etc.

Table 1: Variables of interest

	Control		Hard skils		Soft skills		p-values[a]		
	N	Mean	N	Mean	N	Mean	Hard vs Control	Soft vs Control	Hard vs Soft
Individual demographic characteristics									
Female	1199	0.34	1619	0.36	1596	0.36	0.211	0.222	0.978
Boarding student	1158	0.74	1562	0.73	1535	0.73	0.786	0.536	0.707
Age (years)	1195	20.04	1610	20	1586	20.01	0.416	0.541	0.83
Family background									
Father education: None	1191	0.12	1608	0.12	1585	0.11	0.626	0.454	0.181
Father education: Some primary	1191	0.17	1608	0.15	1585	0.17	0.182	0.911	0.119
Father education: Completed primary	1191	0.12	1608	0.13	1585	0.12	0.332	0.933	0.339
Father education: Some secondary	1191	0.18	1608	0.18	1585	0.19	0.93	0.633	0.673
Father education: Secondary and above	1191	0.41	1608	0.41	1585	0.41	0.969	0.979	0.989
Father source of income: Manual work	1185	0.65	1608	0.69	1582	0.66	0.032	0.649	0.068
Father source of income: Professional work	1185	0.3	1608	0.27	1582	0.3	0.09	0.902	0.049
Father source of income: Unemployed	1185	0.05	1608	0.04	1582	0.04	0.221	0.164	0.852
Mother source of income: Manual work	1194	0.81	1616	0.81	1587	0.81	0.909	0.976	0.876
Mother source of income: Professional work	1194	0.16	1616	0.16	1587	0.15	0.783	0.64	
Mother source of income: Unemployed	1194	0.03	1616	0.03	1587	0.04	0.756	0.38	
Father is alive	1196	0.76	1613	0.74	1589	0.75	0.314	0.60	
Mother is alive	1190	0.89	1617	0.89	1592	0.88	0.932	0.56	
Type of house: Informal structure	1189	0.18	1610	0.18	1586	0.19	0.971	0.57	
Type of house: Fixed structure	1189	0.82	1610	0.82	1586	0.81	0.971	0.57	
Number of rooms at home	1170	4.51	1589	4.6	1560	4.61	0.248	0.19	
Family owns a business	1196	0.53	1616	0.52	1587	0.51	0.666	0.4	
Subjective Family Wealth	1170	4.73	1581	4.77	1562	4.8	0.581	0.36	
Job experience and entrepenurial characteristics									
Age of work initiation	573	14.55	768	14.57	757	14.55	0.902	0.97	
Months worked in formal employment	460	4.2	598	4.19	595	4.12	0.985	0.64	
Exposed to entrepenurial class	1175	0.37	1585	0.33	1563	0.33	0.022	0.0	
Currently has a small business or income generating activity (IGA)	1193	0.44	1612	0.42	1589	0.44	0.37	0.79	

	Control		Hard skils		Soft skills		p-values		
	N	Mean	N	Mean	N	Mean	Hard vs Control	Soft vs Control	Hard vs Soft
Skills									
Digit Span: % digits correct	1190	0.43	1601	0.41	1586	0.41	0.152	0.117	0.881
Ugandan Certificate of Education: Division 1	1177	0.15	1589	0.14	1568	0.14	0.426	0.481	0.923
Ugandan Certificate of Education: Division 2	1177	0.42	1589	0.39	1568	0.4	0.126	0.436	0.418
Ugandan Certificate of Education: Division 3	1177	0.33	1589	0.35	1568	0.35	0.189	0.402	0.609
Ugandan Certificate of Education: Division 4	1177	0.11	1589	0.12	1568	0.11	0.206	0.473	0.556
Compiled score on Raven's test	1173	5.4	1594	5.46	1566	5.39	0.523	0.862	0.379
Time Preferences: Patience	1157	0.28	1567	0.27	1553	0.28	0.739	0.981	0.699
Willingness to Take Risks	1198	0	1618	-0.04	1596	0.04	0.329	0.214	0.016
Math & Business Knowledge	1199	0.6	1619	0.6	1596	0.59	0.959	0.331	0.268
Attitudes toward Enterprise: Leadership	1192	4.17	1609	4.18	1585	4.2	0.438	0.083	0.299
Attitudes toward Enterprise: Perceived Control	1196	4.3	1614	4.34	1590	4.33	0.022	0.074	0.595
Time Preferences: Delta discount rate	1198	-0.01	1618	-0.01	1596	0.02	0.869	0.489	0.353
Time Preferences: Beta time-inconsistency measure	1198	-0.02	1618	-0.01	1596	0.02	0.89	0.28	0.307
N. Ug. Youth Psychosoc. Adj. Scale: Prosocial Behavior	1194	4.26	1609	4.29	1588	4.31	0.161	0.014	0.254
N. Ug. Youth Psychosoc. Adj. Scale: Anxiety & Depression	1192	2.35	1607	2.39	1583	2.42	0.169	0.016	0.261
N. Ug. Youth Psychosoc. Adj. Scale: Self-confidence	1194	4.61	1603	4.56	1574	4.58	0.015	0.252	0.166
Big 5: Extroversion	1175	2.73	1588	2.74	1562	2.7	0.637	0.403	0.157
Big 5: Emotional Stability	1135	3.86	1531	3.88	1497	3.86	0.493	0.934	0.407
Big 5: Openness to Experience	1153	4.16	1549	4.14	1529	4.13	0.62	0.294	0.548
Big 5: Conscientiousness	1144	3.88	1539	3.9	1519	3.88	0.481	0.881	0.356
Big 5: Agreableness	1150	3.63	1559	3.64	1538	3.61	0.759	0.55	0.327
Teacher/Peer/Friend Acceptance	1174	4.24	1571	4.27	1559	4.27	0.259	0.136	0.694
Future Subjective Personal Wealth	1173	7.95	1580	8.02	1563	7.99	0.283	0.516	0.648
Plans to attend university	1177	0.71	1585	0.7	1571	0.71	0.447	0.851	0.537
Distances									
Distance from school to Hard skills Treatment host school	1199	1.49	1619	1.51	1596	1.5	0.596	0.805	0.759
Distance from school to Hard skills Treatment meeting point	1199	1.31	1619	1.33	1596	1.32	0.502	0.687	0.773
Distance from school to Soft skills Treatment host school	1199	1.21	1619	1.24	1596	1.23	0.417	0.599	0.758
Distance from school to Soft skills Treatment meeting point	1199	1.14	1619	1.17	1596	1.16	0.489	0.66	0.786

Table 2: Balance at baseline between treatment and control

Observed Data - Link to the Causal Model

- The observed data were generated by sampling **3,891** independent times from a data generating system compatible with the structural causal model MF.
- This yields **3,891** i.i.d. copies of the random variable

  O=(W, A, Y)~P0.

- The statistical model M for the set of allowed distributions of the observed data is non-parametric.

Counterfactual Outcome

We are interested in the average treatment effect of training on **1)** whether an individual started a business and **2)** log of total earnings (USD):

$$\Psi F(P_{U,X}) = E_{U,X}(Y_1) - E_{U,X}(Y_0) = P_{U,X}(Y_1 = 1) - P_{U,X}(Y_0 = 1)$$

Where Y_a denotes the counterfactual outcomes, if, possibly contrary to fact, the individual went through training $A = a$.

Identification

$\Psi^F(P_{U,X})$ is identifiable because of the randomisation assumption.

All of the observed association between A & Y is due to the causal effect we are interested in i.e we do not have any unmeasured common cause.

If the randomisation assumption holds then ,

$$E_{U,X}(Y_a|W = w) = E_0(Y|A = a, W = w)$$

This gives us the G-Computation Formula

$$E_{U,X}(Y_a) = \sum_w E_0(Y|A = a, W = w)P_0(W = w)$$

$\Psi^F(P_{U,X})$ $\Psi(P_0)$: "estimand"

Descriptive statistics

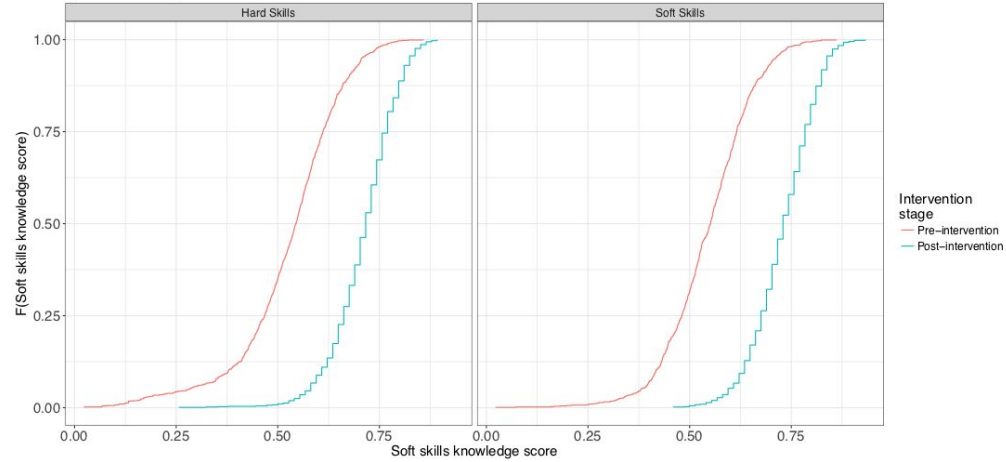|	Control n=1,021		Treatment n=2,870	
	Mean	S.E.	Mean	S.E.
Age	23.5	0.05	23.5	0.03
Female (%)	32.2	1.46	35.7	0.89
Currently in school	0.52	0.02	0.51	0.01
Started a business	0.50	0.02	0.60	0.01
Total earnings (USD-PPP)	5,505	435	5,615	250

Business skills learning

Empirical cumulative density function of Hard skills knowledge score
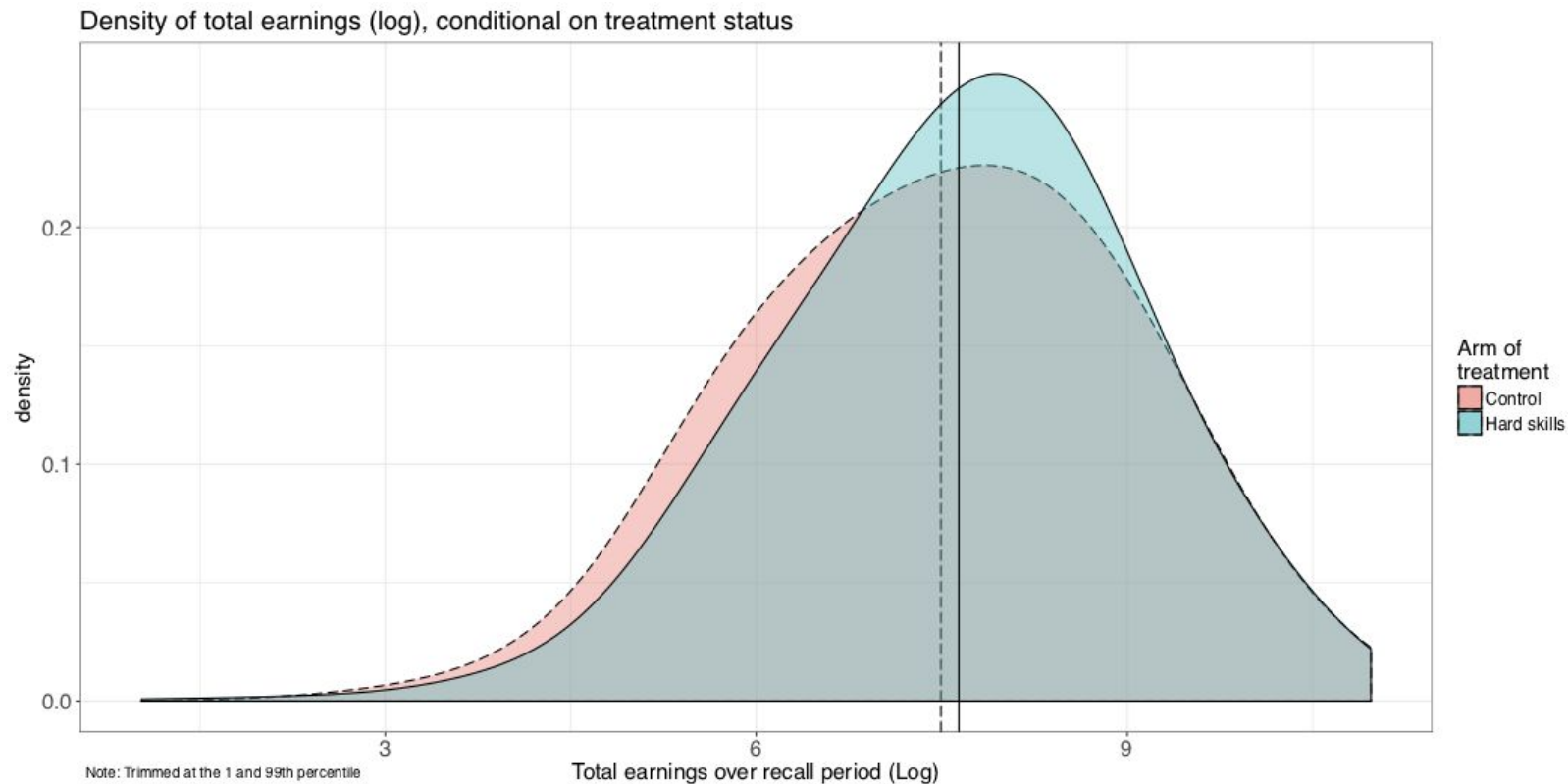By intervention stage
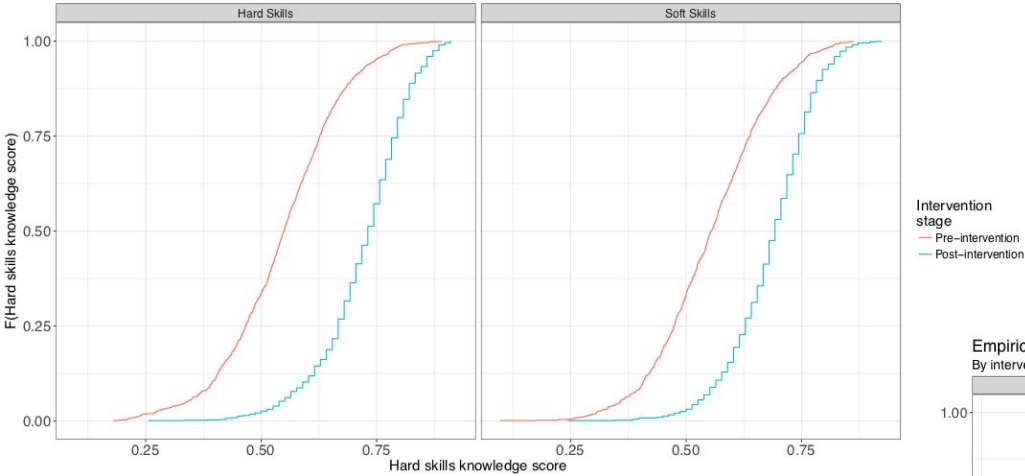
Empirical cumulative density function of Soft skills knowledge score
By intervention stage

Total Earnings on a Logarithmic Scale

Density of total earnings (log), conditional on treatment status

density

Note: Trimmed at the 1 and 99th percentile

Total earnings over recall period (Log)

Arm of treatment
- Control
- Hard skills

Intention-to-treat estimator (lower bound of the ATE)

	GLM	G-computation	IPTW Horvitz-Thompson	TMLE
Ever started a business				
ITT	0.1008	0.1041	0.1006	0.1073
S.E.	0.0732	0.0071	0.0195	0.0174
p-value	3.00E-08	5.15E-26	7.20053E-07	7.13E-10
Log of total earnings				
ITT	0.1369	0.1296	0.1334	0.1399
S.E.	0.0590	0.0475	0.0644	0.0523
p-value	2.03E-02	1.06E-02	4.77E-02	7.48E-03

The SuperLearner library included: glm, glmnet, glm.interaction, and xgboost

Interpretation

- Exposure to entrepreneurial training led to an increase of **0.10 pp** (p-value=7.1×10^{-10}) in the probability of having started a business among a sample of high school graduates in Uganda.
- Relatively, it represents a **21% increase in participation in self-employment activities**.
- Also, the intervention had a positive effect on earnings. On average, training increases total earnings by **13.7%** (p-value=7.5×10^{-3}).

Conclusion

- **Limitations:**
  - Compliance with treatment was not perfect, hence we estimated Intention-to-treat (ITT)
  - May need to include attendance as an additional step in the causal diagram
    - Local Average Treatment Effect (LATE), Complier Average Causal Effect (CACE)
- **Policy: Most training interventions fail, why did this one work**
  - Possibilities include:
    - Selected participants – young recent high school grads
    - Strong intense western/Socratic-based curriculum
    - Teacher selection and training

Conclusion continued...

- **Methods: Use of structural causal models and data-adaptive methods in RCTs can provide increases in efficiency**
  - We don't find significant changes in the point estimates, but precision is improved 1-2 orders of magnitude