Impact Evaluation of End of Program Data from the *Educate*! Randomised Control Trial

Internal Report

2014

Abstract

This report contains a quantitative impact evaluation based on data from a Clustered Randomised Control Trial of a secondary school-level 21st Century Skills, leadership and entrepreneurship education program in Uganda. Using an OLS regression estimator with clustered standard errors, the analysis finds that, out of the twelve outcome variables used to assess the effectiveness of the program of achieving its stated goals, they have reached their target for the outcome variables of business ownership, overall income level, community project ownership, savings behaviour and self-efficacy but not for the indicators of paid-employment prevalence, business or employment income, holding of a school leadership position, business planning, financial literacy or creativity. However, when just females are focused upon, positive impacts on business income and creativity are also observed.

Contents

Introduction

This report summarizes the methodology and findings of an impact evaluation carried out by the US-based NGO *Educate*! which offers an eighteen-month 21st Century Skills, leadership and entrepreneurship education program to secondary school students in an increasing number of schools across Uganda. *Educate*! has implemented a randomized control trial (RCT) to assess the impact of the program on various student-level outcomes related to the program's aims, including livelihood improvement, community participation and practical and 21st Century skills.

Educate!'s RCT was implemented between February 2012 and July 2013. *Educate!* selected 48 schools and randomly assigned these schools to either treatment or control. Treatment schools then received the *Educate!* program, while control schools did not. *Educate!* conducted a baseline survey before the program start and a follow-up survey – hereafter, the end of program survey - immediately after students graduated from the *Educate!* program.

The focal analysis of this report will be on the data collected through the end of program survey with the purpose to identify the impact of the Educate! program on student-level outcomes related to the program's aims for this stage of the program, as outlined in its Logic Model. This analysis will be carried out for the students as a whole group, as well as for females only. The analysis was carried out by the author, a postgraduate student, as part of a research fellowship with *Educate*!. The author has not been involved in the design or implementation of the trial and neither in the data collection through the baseline or end of program surveys.

The value of such an analysis is highlighted by the assertion of Duval-Couetil (2013) who explains that findings from rigorous evaluations can inform internal quality control, as well as aid external accreditation, which Pretorious (2008) asserts is becoming increasingly important due to a rising demand for accountability, especially with regard to leadership and entrepreneurship development programs in formal education institutions and for those that are scaling-up. Furthermore, the rigor of the evaluation heeds the call from a number of scholars who note the weak body of evidence upon which the growing adoption of entrepreneurship education as a development driver is based (Gorman et al, 1997; Glaub and Frese, 2011; Oyugi 2014). In terms of the investigation into the specific impact experienced by women, this is important as the well-documented additional barriers they face in the socio-

economic sphere heightens the need to know what works for them in terms of development assistance.

The report is constructed as follows. The proceeding section will provide context to the analysis, giving an overview of the role entrepreneurship development currently occupies in Uganda, as well as reviewing the theoretical and empirical landscapes. Section 2 will outline details of the program itself as well as of the trial design, followed by an overview of the analytic methodology in Section 3. Sections 4, 5 and 6 will then respectively present the results of balance testing and the model specifications, the results of the analysis, and a discussion of these results, mainly regarding their relevance to *Educate*! and to the field of entrepreneurship education. This will then be rounded-off by some concluding remarks which will contain suggestions for further research as well as the weaknesses and areas for possible improvement of the study.

Section 1 - Background

1.1 Entrepreneurship Development as a Solution to Youth Unemployment in Uganda

In 2002, the Ugandan government announced that entrepreneurship would be integrated as an official subject in Ugandan secondary schools, primarily in a bid to curtail youth unemployment, one of the many economic and social issues afflicting the country (Wiegratz, 2009). However by 2012, Ugandan President Yoweri Museveni was still describing this issue as "out of control" (Edyegu, 2012). In Uganda young people represent about half of the unemployed, while those employed work in jobs in the informal economy, characterized by low income and unstable conditions (ILO, 2014). Accordingly, young people are hampered to develop their full economic potential, while the state suffers of multiple costs inherited by the huge share of youth unemployment with James-Wilson (2008) describing it as a "profoundly destabilising force" (P.2). In addition to the tax base issues which accompany mass unemployment amongst any group (Nafukho & Muyia, 2009), DeJaeghere & Baxter (2014) outline that unemployment amongst the young can contribute to deterioration in national levels of health, crime, civil conflict and social unrest, especially in developing countries.

In addition to its adoption in the national curriculum, the past decade has also seen a rise in international entrepreneurship development NGO's extending their operations into the country, partly as a consequence of the intervention's ineffective adoption in public policy (ibid). Along with *Educate*!, these organisations include Junior Achievement (See Youth Unemployment Network, 2013) and the Start and Improve Your Own Business program of the ILO (See Mathisen et al, 2011).

This rise of entrepreneurship development as a remedy to youth unemployment in Uganda is a reflection both of similar moves across Sub-Saharan Africa, in countries including Kenya, Botswana and South Africa (See Nelson & Johnsen, 1997; Farsad, 2002; Orford et al, 2004), as well as of a widespread view held amongst the international community (See Bhargava, 2007). Awogbenle and Iwuamadi (2010) explain that youth unemployment and a country's economic growth are inextricably linked, and most of the suggested mechanisms through which entrepreneurial activity is deemed to decrease youth unemployment come indirectly through boosting economic growth. These include the view that entrepreneurs act as facilitators of the division of labour - a perceived key to economic progress dating back to the work of Adam Smith (Michael, 2008) - and that they increase the efficiency of an economy

through the identification and exploitation of profit opportunities (Bhargava, 2007), which also fuels innovation through Shumpeterian creative destruction (Oosterbeek et al, 2010). An indirect economic growth facilitator is also suggested to be provided by social entrepreneurs, who are characterised as applying the innovative tendencies of entrepreneurs to address growth-hindering social issues (Mair & Marti, 2004; Seelos & Mair, 2005). With these mechanisms in mind, Nafukho and Muyia (2009) go as far as to state that "entrepreneurship has been the driving force behind every nation's economic development" (P.96).

In terms of its direct links to decreasing youth unemployment, the main suggested mechanism is through entrepreneurs being creators, rather than occupiers, of employment, which grows the job market whilst also allowing the circumvention of barriers faced by the young to entry and progress in the wage employment sector such as an emphasis on prior work experience (Bhargava, 2007; Subba Rao & Durga Prasad, 2007). It is also believed that entrepreneurship development as an intervention is more suited to the young, who are seen as being more responsive to new economic opportunities (Awogbenle and Iwuamadi, 2010); have been found to possess greater levels of satisfaction in self-employment than other age groups (Blanchflower and Oswald, 1998); are easiest to reach through educational channels with more malleable minds (Mahlberg, 1996; World Economic Forum, 2009); and are the key to creating a sustained "culture of entrepreneurship" (Nelson and Johnson , 1997).

In terms of the specific potential of entrepreneurship development for addressing unemployment amongst young women, de Mel et al. (2012) explain that self-employment makes up a large portion of women's economic activity in developing countries and is an important source of their empowerment.

1.2 Review of Existing Entrepreneurship Development Impact Evaluations

The field of entrepreneurship development is deemed to suffer from a dearth of rigorous analytical methodology utilisation for evaluating program effectiveness (See Gorman et al 1997; Henry et al, 2005; Matlay, 2005). This issue is highlighted by the systematic reviews of both Harper and Finnegan (1998) and Glaub and Frese (2011) who assess all available studies carried out in developing countries. Only a very small percentage deemed not to suffer from serious methodological issues based around sample size, length of study – an issue particularly relevant to the young who may experience delayed impact (Mwasalwiba, 2010) -, rigor, internal validity of design and measures and the quality of data analysis utilised. Uniformity of evaluation is also suggested to be greatly hindered by the heterogeneity of

program design, target audience and target outcomes, as well as success indicators used by researchers (McKenzie and Woodruff, 2013).

The majority of available studies on entrepreneurship education are set in the context of developed countries and their findings are mixed. For instance, a RCT on a voluntary entrepreneurship education program in the US found that business ownership and employment were improved in the short-term but that these benefits disappeared in the long-term (Fairlie et al, 2012). In a study by Oosterbeek et al (2010) on an entrepreneurial vocational scheme at a college in the Netherlands no statistical significant impact could be identified on the entrepreneurial skills of participants. However, the majority of the findings of developed country studies have been more positive, finding significant impacts on outcomes such as achievement and internal locus of control of recipients (Hansemark, 1998) and on self-efficacy (Ehrlich et al, 2000) both in the US, as well as on the desirability and feasibility of entrepreneurship for graduates of a program in Australia (Peterman & Kennedy, 2003), an on general entrepreneurial capabilities in the UK (Kolveried & Moen, 1997; Pittaway & Cope, 2006).

Although insightful, Martinez et al (2010) explain that context is hugely important to entrepreneurship education outcomes, and findings from developed country studies may not be applicable to countries of the global South. This is attributed to differing types of entrepreneur intention and motivation, skill level and opportunities as well as cultural attitudes towards entrepreneurship, market environment, schooling quality and availability and government support, all of which can influence the nature and impact of entrepreneur education programs.

Shifting the focus to studies carried out in developing countries, Table 1 below outlines the most relevant in terms of context within which they are delivered, program type and by the rigor of the evaluation methodology.

Table 1:

Authors	Country	Program	Methodology	Findings
Orford et al. (2004)	South Africa	Entrepreneurship education curriculum in secondary schools	• Cluster Randomised Control Trial • Randomisation at school level	*Entrepreneurial Skills* • Increased level of understanding of basic financial concepts and principles *Soft Skills* • Increased level of confidence in ability to start a business and on achievement orientation
Klinger & Schündeln (2007)	El Salvador, Guatemala & Nicaragua	Voluntary training on technical business skills and core entrepreneurial behaviours for potential entrepreneurs	• Quasi-experimental regression discontinuity design • Randomisation around entrepreneurial potential score threshold. • Data collected one year after program end.	*Employment Outcomes* • Increased level of self-employment No other outcomes were assessed.
Karlan & Valdivia (2011)	Peru	Entrepreneurship training provided to female microfinance recipients	• Cluster Randomised Control Trial • Randomisation at the lending group level • Data collected between one and two year after program end	*Entrepreneurial Skills* • Improvement in business knowledge

Bandiera et al (2012)	Uganda	Two-pronged, voluntary adolescent (age 14-20) female empowerment program aiming to teach entrepreneurial and life skills.	• Cluster Randomised Control Trial • Randomisation at the lending group level • Data collected two years after program start	*Employment Outcomes* • Increased level in labor force participation, primarily self-employment *Entrepreneurial Skills* • Higher level of entrepreneurship skills
World Bank (2012)	Liberia	1 Year long program for 15-24 year old females, consisting of 6 months of classroom training and 6 months of work placement	• Randomised pipeline trial where participants were randomly allocated treatment at one of two rounds. • 1273 were assigned to round 1 and 769 to round 2. • Data collected immediately after the classroom phase	*Employment Outcomes* • Increase in employment for participants • Increase in income level *Entrepreneurial Skills* • Higher level of savings
International Youth Foundation (2013)	Kenya	Eight-week comprehensive employability program for women comprising training on technical skills, entrepreneurship skills and life skills	• Randomised Control Trial • Randomised at the individual level • Data collected 6 months after program end	*Employment Outcomes* • Higher access to quality employment *Soft Skills* • Increased level of job skills • Increased level of self-confidence

Other than the Orford et al study mentioned in the above table, there does not seem to be any rigorous studies carried out on entrepreneurship education at the secondary school level in a developing country setting. The only other study that the author is aware of is that of Mathisen et al (2011) who carried out an assessment of the Ugandan branch of Junior Achievement, an after-school entrepreneur education program aimed at 16-18 year olds. This involved interviews with key informants and group sessions with program students directly after completing the 6-month course. From these interviews, the study found that students demonstrated improvements in leadership and motivation thanks to the program.

Although by no means unequivocal, one may take a reasonably positive view of the impact of entrepreneurship education from previous studies in developing countries. As well as findings summarized in Table 1 indicating at least one element of positive impact, the aforementioned review by Glaub and Frese (2011) found that the most rigorous studies report positive effects on targeted psychological traits, business management skills and business performance.

Section 2- Program and Trial Design

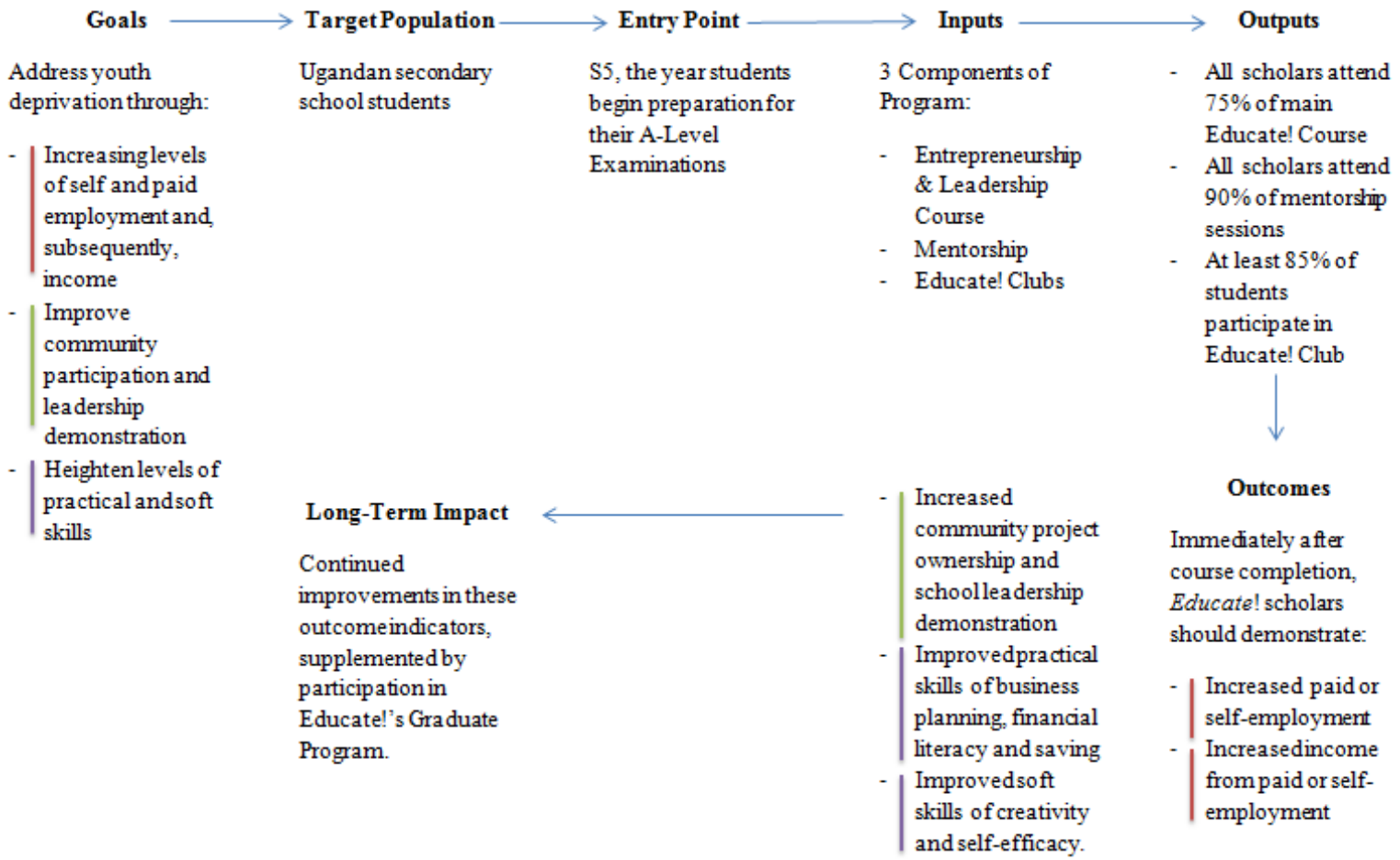
2.1 The *Educate*! Program

In the past *Educate*! has been implemented the program in 54 schools across Uganda. Since 2014 *Educate!* has engaged in a major scale up process and now operates in more than 200 secondary schools in Uganda. Their long-term plan is to see the program fully integrated into the Ugandan secondary school curriculum and also to spread to other countries in Sub-Saharan Africa.

Since 2008, *Educate*! has administered a voluntary in-school program for selected 16-18 year olds, implemented for eighteen months during the final two years of secondary school (S5 and S6). The main aim of this program is to support secondary students in their transition between school and work through transforming them into successful entrepreneurs, employees, leaders and social activists. All students in their penultimate year of secondary school are eligible to apply to the *Educate!* program and are then subjected to a participation eligibility test that assesses their entrepreneurship and leadership propensity. Based on this selection process, a maximum of 45 and a minimum of 30 students are selected from each school.

As outlined in the organisation's Logic Model below, the primary input of the program are weekly lessons on Entrepreneurship and Leadership, complimented by mentoring sessions and Student Business Clubs activities. The *Educate*! program emphasises experience-based learning through practically designed lessons, building experience in developing business plans and managing businesses through club activities and learning vocational skills for potential businesses, such as liquid soap making and paper recycling. Lessons are administered by a specifically trained *Educate*! Mentor. The *Educate*! Mentor also supports the club activities, though these are primarily led by the students themselves. *Educate*! mentors are tasked with acting as a teacher, friend and role model to the students and are recruited as recent University graduates.

Figure 1: *Educate*! Logic Model



2.2 The *Educate*! Clustered Randomised Control Trial

i. Background on Randomised Control Trials

A raft of scholars extoll the virtues of Randomised Control Trials (RCTs) as the most rigorous impact evaluation tool, primarily due to its use of chance to create treatment and comparison groups for causal inference (See Duflo et al, 2006; Imbens & Wooldridge, 2008; Barahona, 2010; Khandker et al, 2010). These scholars explain that using anything other than chance threatens the requisite unconfoundedness assumption that requires no variable which could shape potential outcomes to influence treatment inclusion.

The primary source of the violation of this assumption is through selection bias, whereby either those implementing the program are required to select individuals or the individuals are required to select themselves into the program. This is deemed to open up the possibility of participants' characteristics systematically affecting their selection into the program and, consequently, compounding their comparability with individuals who do not participate in the program. It also becomes more difficult to find individuals who do not participate in the program but who are representative of those selected to effectively reflect the performance in the absence of the program. It is asserted that this issue is further challenged due to the extreme difficulty in measuring and controlling for such bias, with the confounding characteristics of treatment participants often being unobservable.

As aforementioned, a lack of rigor - either through the absence of a comparison group or through their inadequate construction - is suggested to be a major issue in the field of entrepreneurship education assessment. Martinez et al (2010) explain that, with specific regard to entrepreneurship education programs - particularly those which are voluntary and selective - selection bias is an particularly acute issue due to motivation and ability dictating elements of entrepreneurial success.

ii. Focal Outcomes

The *Educate*! trial was designed to assess a number of indicators that aim to represent the effective achievement of the program's outcomes. As outlined in Figure 1, by the time that students graduate from the program, *Educate!* expects improvements in indicators of all three major outcomes, which are presented in Table 2 below. As no target size of improvement is set by *Educate*!, the effectiveness of the program in meeting its outcome targets will be

judged as any statistically significantly higher level of these indicators for the treatment group compared to the control group.

Table 2:

Outcome	Indicators
Increased levels of self-employment and paid employment and, subsequently, income	• Business ownership prevalence and business income • Paid employment prevalence and wages • Overall income level across all students
Increased youth community participation and school leadership demonstration	• Prevalence of community project ownership • Prevalence of leadership roles held at school
Improved practical and soft skills	• Practical Skills: Higher composite scores for business planning, financial literacy and savings behaviour • Soft Skills: Higher composite scores for creativity and self-efficacy

With regard to the composite scores used to measure business planning, financial literacy, savings behaviour, creativity and self-efficacy, these metrics were specifically designed by *Educate*! and Rachel Steinacher through intensive qualitative research to adequately measure these skills in the Ugandan context (for full list of the questions used for each metric, see Appendix 4).

iii. Trial Design and Implementation

The particular trial design implemented by *Educate* is referred to as Clustered RCT, as randomisation was carried out at the school level. Torgerson and Torgerson (2001) state that this is a particularly common method used in educational research as it avoids the strong threat of contamination of the treatment group when there are members of the treatment and control group sharing the same school or class.

For the trial, of the 111 districts in Uganda, six districts were purposively selected by *Educate!*. District selection was initially based on the criteria to have at least 8 A-level (upper secondary) schools with more than 100 S5 students and more than 40 students in the entrepreneurship track. Yet, these criteria turned out to be too restrictive in the Ugandan

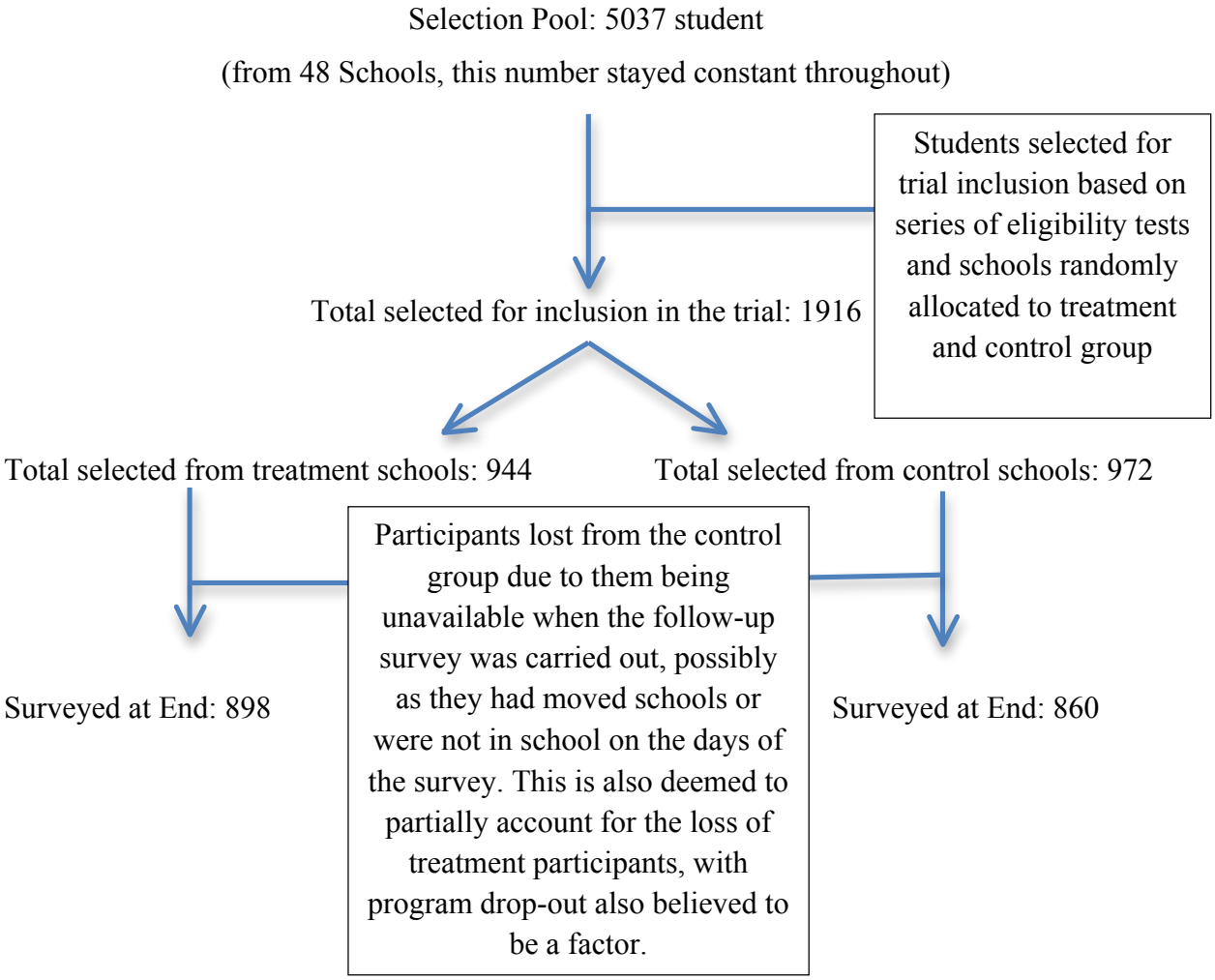
context. Consequently, the selection criteria have been relaxed to have merely more than 8 A-level schools with more than 40 S5 students (first year of upper secondary). Eligible districts have been chosen with preference towards populous districts to account for potential school drop-outs as well as towards a fair regional distribution. From each of the chosen districts, 8 schools that consented to inclusion in the trial were then selected to be included in the RCT by *Educate*!. Accordingly, a total of 48 schools were recruited to participate in the trial (see Appendix 1 for a list of districts and schools).

Unblinded IPA field staff were then required to select 30-45 students in treatment and control schools. Selection was based on a short survey to all S5 students to ascertain their interest in participating in a leadership and entrepreneurship course and if they had previous leadership and/or entrepreneurial experience as well as on games assessing the students literacy level as well their cognitive ability. Students were assigned a score based on the survey and game participation and the top 45 students in all schools were selected to be the treatment/control samples. The purpose for conducting this process in both the treatment and control schools was to maximize the chance that the treatment and control participants are comparable. After students were selected, their consent was sought for inclusion in the trial.

After student selection, the program was then randomly allocated to 24 schools and the other 24 were allocated to the control group, with allocation being stratified by district. Randomization has been carried out by an external party, namely the Principal Investigator Nathan Fiala, in partnership with Innovations for Poverty Action (IPA). In each district for which 8 schools had been selected, 4 schools have been randomly selected to receive the *Educate*! program, while the remaining 4 were selected to act as control schools.

A baseline survey was carried out before the program began in 2012 by IPA and included all students that had consented to inclusion. A follow-up survey – the end of program survey – was administered face-to-face by field enumerators hired by *Educate!* immediately after students graduated in 2013 and included only those who had completed the program and those still in the control group cohort (See Appendix 3 for the full survey). As indicated in Figure 2 below, 95% of the treatment students and 88% of the control students in the baseline survey were reached and surveyed for the first follow-up survey. The data was entered by an external data-entry company.

Figure 2: Participant Flow Diagram



Source: Author's own

iv. Trial Power.

The power of an experiment – here the RCT - is a measure in the ability to detect an effect of the program if such an effect does exist. According to McKenzie and Woodruff (2013), "in order to obtain credible and useful estimates, studies must have sufficient power" (P.8). The main determinants of the power of an experiment are the size and heterogeneity of the sample

The primary means of assessing the power of a model for RCTs is to calculate the Minimum Detectable Effect Size (MDES), which represents the smallest treatment effect that can be identified through the analysis of the trial data, with any impact below the MDES not being detectable (Jacob & Zhu, 2009).

The calculation used in order to calculate the MDES for clustered RCTs is outlined below. This differs from the standard method as it includes the Intra-Cluster Correlation (ICC) coefficient that represents the cluster-level variation as a proportion of the total variation in

the data and, as it increases, the statistical power of a trial decreases as individuals within a cluster offer less unique information (Cameron and Miller, 2013).

$$\text{MDES} = M_{J\text{-}K} \sqrt{\frac{\rho(1 - R^2_C)}{P(1 - P)J} + \frac{(1 - \rho)(1 - R^2_I)}{P(1 - P)nJ}} \quad (1)$$

(Taken from Hutchison & Styles, 2010)

Where $M_{J\text{-}K}$= a multiplier based on the t-distribution, ρ= the ICC, R^2c= the increase in the R^2 statistic due to inclusion of cluster-level covariates, R^2i=the increase in the R^2 statistic due to inclusion of individual-level covariates, P=the proportion of the sample allocated to the treatment group, *n*= the number of individuals per cluster and *J*= the total number of clusters.

As there are a number of outcome variables, this calculation was carried out with the average values taken from all of the models, with the average MDES for the trial being given as 0.228 (See Appendix 2 for further calculation details). This means that the model would not be able to detect an effect of the program that is less than 0.228 standard deviations away from the baseline mean of any of the outcome variables. In terms of the subgroup analysis of females, due to the decreased sample size, this figure rises to 0.326.

Although leading scholars such as Jacob Cohen and Mark Lipsey respectively suggest MDES values of 0.2 and 0.15 as representing a strong model (Hutchison & Styles, 2010), the favorability of MDES values is widely believed to be context specific (Jacob & Zhu, 2009). In this case, with the comparison of a full program against no treatment at all, one may assert that the MDES does not need to be extremely precise as, if the program is effective, impact would not be subtle.

v. Survey Design

As indicated earlier, data has been collected through a baseline survey before the program began and through an end of program survey immediately after students graduated from the *Educate*! program.

The purpose of the baseline survey was mainly to provide information on general demographic characteristics of the selected students as well as on their economic activities and civic engagement. The end of program survey included questions about the participants' economic activities, their civic engagement, as well as on their personality and attitudes. For

personality and attitudes, *Educate!* and Rachel Steinacher specifically developed the before-mentioned composite metric scores to measure students' practical and soft skills. Yet, since those were only developed after the implementation of the baseline data collection, the non-availability of these variables in the baseline data set rendered it impossible to measure their balance across treatment and control students and to include their baseline values in the regression model to increase their estimation precision.

Furthermore, measuring income for business owner has proved to be difficult with the design of the end of program survey. As stated by McKenzie and Woodruff (2013), collecting data on profits and revenues poses certain problems, as owners of the smallest businesses usually do not keep records (p. 11). The question to measure income of business owners included in the end of program survey has been the following:

How often do you earn income with your business and how much do you earn?

☐ I get monthly. In an average month I make _____UGX.

☐ I earn money when my plants produce. I usually make _____UGX at the end of the growing season.

☐ I have shares in a project so I get money once or twice a year. I usually get _____UGX from my shares.

☐ I earn money when I am able to make a sale, which happens from time to time. I usually make _____ (number of sales) in 1 month. I usually make _____UGX per sale.

It becomes obvious that only the first two answer options have a clear time dimension attached to them, while the third answer option on shares does not. That is why, all responses falling into this category had to be omitted from the income measurement. It has also been found that the fourth answer option on sales has been confusing for respondents, revealing unrealistic data. Furthermore acknowledging that sales do not directly correspond to income for business owners, responses falling into this category had to be omitted. Due to these challenges the sample size for business income decreased.

vi. Potential Issues with the Trial Design

As stated by Lewis and Warlow (2004), "just because a study is randomised does not mean it is unbiased" (P.181), with a number of potential sources of bias arising from various elements of a randomised trial's design such as a small sample size. Puffer et al (2003) add that, in the case of clustered RCTs, such bias can occur both at the cluster and individual level, and a few of their suggested potentially problematic design elements can be identified in the *Educate*! RCT.

For the student selection, the IPA field staff who were tasked with carrying out the student selection process after school allocation to treatment and control may have biased the student selection, eventually through carrying out of the selection process with less rigour in control schools as they knew those selected would not actually be receiving the program, although the direction of this possible bias is not clear. In Section 4 of this report, balance testing will be carried out, partly to assess whether any of these potential sources of bias materialised.

Finally, the fact that data was not collected from students who dropped out of the trial, provides further scope for bias through differential attrition. In fact, if there would be any systematic difference between drop-outs among treatment students and control students this would lead to biased results. Accordingly, it has been investigated separately if such systematic difference exists and subsequently, the analysis has been using treatment-on-the-treated estimates rather than intent-to-treat estimates, as explained further in the next section.

Section 3 - Analytic Methodology

3.1 Analysis of the Treatment on the Treated

Dunning and Hyde (2008) explain that analysis of the effect of the treatment on the treated (TOT) involves assessing just those who participate in the program under evaluation with those in the control group. This is opposed to the other main analytical method of assessing impact based on intention-to-treat (ITT) that includes all those initially selected for inclusion in the trial, regardless of their participation. The latter option yields more conservative results as it captures the effect of *offering* training irrespective of actual take-up and thereby avoids incurring bias caused by differential attrition.

Since *Educate!* did not collect data from students that dropped out, in this analysis the TOT method will be applied. Accordingly, the estimates will indicate the average impact of the program for students who fully complied with the treatment. As mentioned earlier, around 8% of students have been included in the baseline survey but not in the end of program survey with a least a fraction of these being due to trial drop-out, the size of which is not known. As indicated earlier, it will be investigated if this share of drop-out was systematically different across treatment and control students.

3.2 Accounting for the Clustered Structure of the Data

For the data analysis on student-level outcomes in a clustered dataset, it is important to adjust the analysis for clustering. In the case of school-level clustering – the case at hand here - outcomes of students within the same school can be correlated because of the exposure to the same teacher or the facilities of the school, for example (Hutchison & Styles, 2010). It becomes apparent that for regression analysis clustered data provides less unique variation in the data and also leads to a correlation of the error term by cluster, as individuals within a cluster have their outcomes similarly affected by such cluster-level variables. As a result, if clustering is not accounted for, most standard estimation models - such as OLS - produce underestimated standard errors. Underestimated standard errors are biasing the significance testing with misleadingly narrow confidence intervals and low p-values, inducing someone to falsely interpret a p-value that undercuts the 5 percent significance level as a proof of an effect (Cameron and Miller, 2013).

It is possible to account for such data structure by either applying multi-level modelling or applying clustered standard errors. The former accounts for the clustering by separating the error term into an idiosyncratic fraction that varies across individuals and another cluster-specific fraction, which is either treated as fixed or random across all clusters. Alternatively, it is possible to apply clustered standard errors to the standard estimation model simply adjusting for the introduced bias of less unique information through clustering. It is widely regarded that the multi-level modelling approach provides more accurate estimation results over the approach of clustered standard errors (Wooldridge, 2002). Yet, these desirable properties hold only under very strong assumptions, which were tested and found violated in the context here. Most importantly, multi-level modelling assumes a normal distribution of the error terms. The Shapiro-Wilk test has shown that this assumption was violated for a number of outcome variables of this trial.

Therefore, this analysis accounts for clustering by applying clustered standard errors to the estimation model. Assumptions of this model have also been tested and found to not be violated. At first, the model assumes no correlation across clusters and analysing the correlation matrixes of the focal outcome variables and schools, no such correlation was found. The model also requires a large amount of clusters – ideally around 50 - for valid and reliable inference and with this study including 48 schools, it seems possible to conclude that this model is the Best Linear Unbiased Estimator for this analysis.

3.3 Accounting for Attrition Bias

In the *Educate*! trial there was some potential for attrition bias as data had not been collected from students who dropped out of the program between the baseline survey and the end of program survey. Any systematic difference between these dropouts would imply biased impact results.

Hence, an analysis has been conducted on the baseline characteristics of these drop-outs. It was found that no such systematic differences prevail among the drop-outs across the groups, with only little statistically differences among the key characteristics that could potentially influence outcomes. Accordingly, there does not seem to be differential drives of attrition between treatment and control students and the treatment-on-the-treated estimates will yield unbiased results.

3.4 Selection of Control Variables

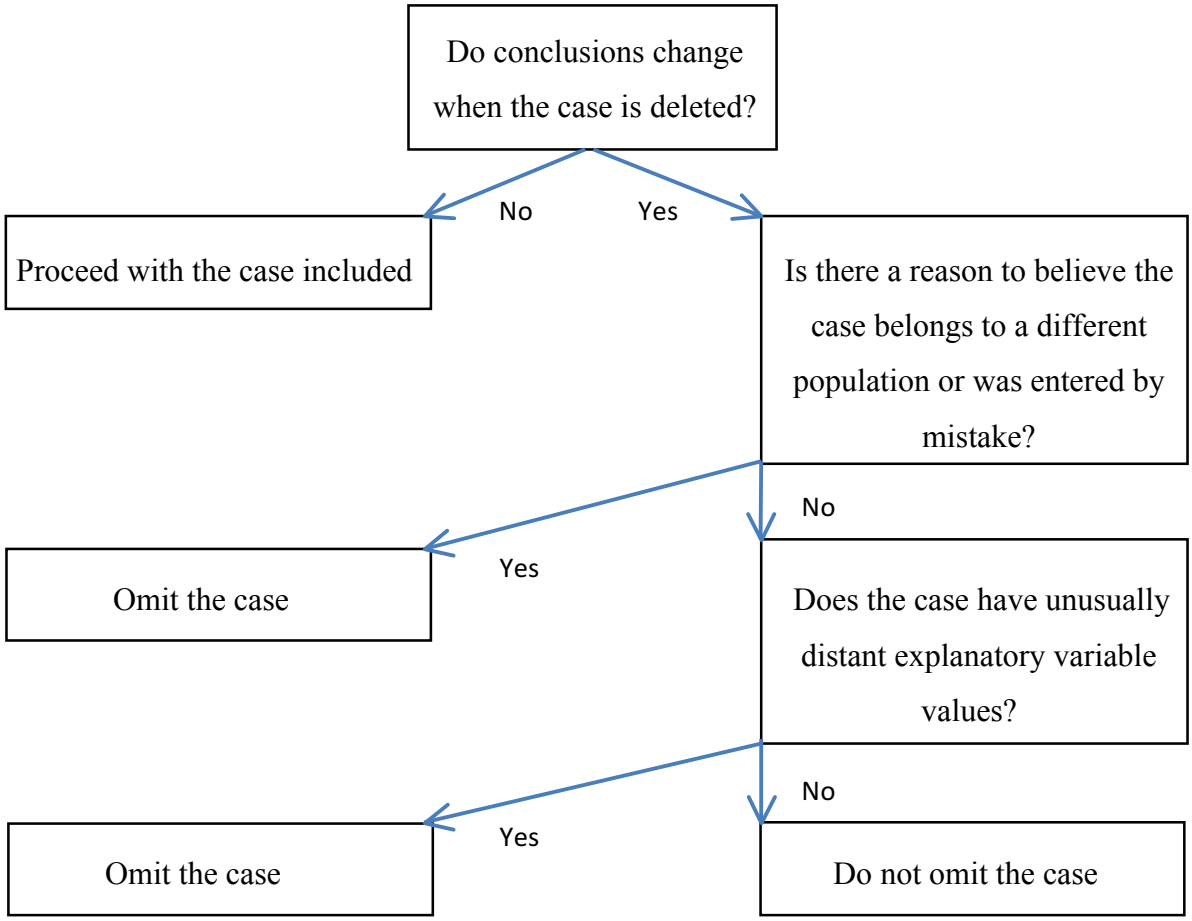
According to Imbens and Wooldridge (2008), the appropriate use of control variables can greatly increase the precision of a statistical model. Most importantly, an analysis of RCT data is only valid if treatment and control groups are perfectly balanced. Though randomization is deemed to be the most effective means of creating balanced comparison groups, it is still possible that treatment and control groups inherit statistically significant differences with regards to certain characteristics. Accordingly, balance testing will be carried out and inform the choice of control variables. Altman (1985) explains that both the size of the imbalance as well as the perceived effect on the outcome variable should dictate the inclusion of a variable into the regression model, emphasizing that significance testing is irrelevant as this merely assesses the randomness of a difference.

In order to avoid what is termed "specification searching" by Duflo et al (2006, p.66) assessing the correlation of potential regressors with the outcome variables in order to decide which to include, a set of variables deemed to have the strongest influence were chosen for each outcome along with any other relevant variables which were found to be highly imbalanced between the treatment and control groups. The full models are laid out in Section 4.2.

3.5 Treatment of Outliers and Missing Values

For each of the continuous outcome variables, extreme outliers that had a highly distortionary effect on mean values were removed before the analysis. The identification of outliers was carried out using the strategy outlined below, based on the advocacy of O'Halloran (1996).

Figure 3: Outlier Removal Strategy



In terms of missing values, these were dealt with on a variable-specific basis. The most relevant guidelines are mentioned here, while a full summary of the approach is provided in Appendix 5.

For questions that require an initial yes/no statement followed by a list of related questions that are to be answered if 'yes' was answered to the initial question, the following rules have been applied:

- If the respondent answered 'yes' but then did not answer any of the follow-up questions, the initial yes/no statement has been replaced with a missing value.
- If the respondent answered 'no' or left the question blank but then did answer all of the follow-up questions, the answer to the initial question has been changed to 'yes'.

This applies to questions on business or community project ownership, for example.

For questions on income, all missing values have been replaced with zeros, assuming missing values correspond to no income.

For the composite metric scores assessing practical or soft skills, missing values have been assigned to students that have answered less than a third of the composite questions.

Section 4 - Balance Testing and Model Specification

4.1 Balance Testing

i. Balance Data for the Baseline Cohort

As noted in Section 2.2vi, imbalance in RCTs can occur at both the cluster and the individual level. Accordingly, Tables 4a and 4b below respectively present the school level and the student level balance between variables that could affect outcomes for the treatment and control groups. It also presents the percentage that the mean differs between the two groups as a percentage of the treatment group mean and the p-value given by the two-tailed t-test. Although a large number of variables were tested, only an excerpt is shown here.

Table 4a: Balance Data for School-level characteristics

Variable	Treatment	Control	Difference	P-Value
Students taught:				
Average No. S5 Students Ave No.	174	160	+7.8	0.774
Gender taught (%):				
All Girls	12.5	4.2	+8.3	0.460
All Boys	8.3	4.2	+4.1	
Mixed	79.2	91.7	-12.5	
Ownership (%):				
Public	37.5	34.8	-2.7%	0.846
Private	62.5	65.2	-2.7%	
Income Level:				
Average Termly Tuition Fee (in UGX)	386,115	374,506	+7.0%	0.902

Significance at 90% indicated by (*); at 95% by (**) and at 99% by (***)

As it becomes obvious from Table 4a, randomization at school level was relatively successful. A higher percentage of single-sex schools as well as public schools were found in the treatment group. Yet, none of these differences is statistically significant across the groups.

Table 4b: Balance Data for Student-level characteristics (baseline cohort)

Variable	Treatment	Control	Difference (%)	P-Value
Personal Characteristics				
Age	18	18	+0.07	0.332
Female (%)	46.7	40.9	+5.8	0.011**
Participant Economic Well-being				
Wealth Scores	2.6	2.6	-0.0	0.9867
Participant Education				
O-Level Grade – Division 1	29.7	28.1	+1.6	
O-Level Grade – Division 2	43.7	42.4	+1.3	
O-Level Grade – Division 3	22.8	25.4	-2.6	0.569
O-Level Grade – Division 4	3.7	3.9	-0.2	
O-Level Grade – Division 5	0.1	0.3	-0.2	
Studied Entrepreneurship Subject (%)	32.9	33.5	-0.6	0.424
Economic Activity				
Likely to start business in the future (%)	81.5	77.0	+4.5	0.079*
Own a business (%)	41.5	44.5	-3.0	0.184
Business income (in UGX)	275,563	189,576		0.6147
Have a paid job (%)	26.8	31.6	-4.8	0.022**
Wage income (in UGX)	56,985	81,532	-43.1%	0.0639*
Leadership				
Have Community Project (%)	8.1	7.6	+0.5	0.701
Hold Leadership Position (%)	63.1	61.4	+1.7	0.439

Significance at 90% indicated by (*); at 95% by (**) and at 99% by (***)

At student level, randomization equally yielded relatively balanced group of treatment students and control students. It was found that in the treatment group a higher share of female students has been selected. Students in the treatment group also indicated a higher ambition of starting a business in the future compared to the students in the control group.

ii. Balance Data for the End of Program Cohort

Although assessing balance between students of the baseline cohort is useful for illustrating the neutrality of the randomisation process, this is inadequate for aiding the choice of control variables to be used to maximise the precision of the statistical analysis due to the possibility of differential attrition between the baseline and end of program stages. Therefore the table below displays the balance data across students for the end of program cohort.

Table 5: Balance Data for Student-level characteristics (end of program cohort)

Variable	Treatment	Control	Difference (%)	P-Value
Personal Characteristics				
Age	19	19	-0.12	0.032**
Female (%)	48.6	40.2	+8.4	0.001***
Participant Economic Well-being				
Wealth Scores	2.6	2.7	-0.1	0.9942
Participant Education				
O-Level Grade – Division 1	30.3	29.4	+0.9	
O-Level Grade – Division 2	43.7	42.2	+1.5	
O-Level Grade – Division 3	22.0	25.1	-3.1	0.407
O-Level Grade – Division 4	3.8	2.9	+0.9	
O-Level Grade – Division 5	0.1	0.4	-0.3	
Studied Entrepreneurship Subject (%)	26.678%	28.175%	5.61%	0.8008
Economic Activity				
Likely to start business in the future (%)	80.9	77.4	+3.5	0.85
Own a business (%)	40.4	44.7	-4.3	0.076*
Business income (in UGX)	304,810	169,215	+44,5%	0.498
Have a paid job (%)	26.8	32.1	-5.3	0.018**
Wage income (in UGX)	55,385	81,225	-46.7	0.0773*
Leadership				
Have Community Project (%)	7.5	7.3	+0.2	0.911
Hold Leadership Position (%)	64.4	61.5	+2.9	0.250

Significance at 90% indicated by (*); at 95% by (**) and at 99% by (***)

This provided the main basis for control variable choice in the model specifications outlined in the following section.

4.2 Model Specifications

Following the balance testing, all estimation models will control for gender acknowledging the imbalance in the dataset as well as for the wealth score as proxy of the socio-economic background of the students considering its possibly high influence on the targeted student-level outcomes. All specifications will also include a control for the baseline values of business ownership, community project ownership and wage employment considering the identified imbalances as well as the high relevance. Further selection of control variable will be specific to the models for each outcome.

i. Increased levels of self-employment and paid employment and, subsequently, income

To assess this outcome, the impact of the program will be identified on business ownership and income, wage employment and wages and overall income across students.

In all of these five model specifications, we will control for the score on the O-level certificate as a measure for students' educational background and the score on the financial literacy test as proxy for pre-existing entrepreneurial knowledge.

For all three model specifications related to income – business income, wage income or income across all students – we will also control for the baseline value of those.

ii. Increased youth community participation and school leadership demonstration

This outcome will be assessed against two models determining the impact on community project ownership and leadership demonstration as a result of participation in the *Educate*! program.

In these two model specifications, we will control for the score on the O-level certificate as a measure for students' educational background but not for the score on the financial literacy test, as these specifications inform the impact of the program on leadership behaviour. The model specifications will also control for the baseline values of the respective outcomes, namely community project ownership and leadership role occupation.

iii. Improved practical and soft skills

To assess the impact of the program on practical and soft skills, the relationship between the program and business planning, financial literacy and savings as well as on creativity and self-efficacy will be investigated.

In the specifications related to practical skills, controls for the score on the financial literacy test were incorporated as proxy for pre-existing entrepreneurial knowledge. In the specifications on soft skills, baseline values of the assessed soft skills were included in the model where measures were available, such as for self-efficacy.

Section 5 - Results

5.1 Results from descriptive statistics

It is common practice in reporting results of RCTs to display both the outcome of comparisons between the treatment and control groups without controlling for the influence of other variables, as well as a model using relevant regressors, as specified in section 4c. Tables 6a – for the entire group of students - and 6b – for female students only - present the results of the former.

Columns (2) shows the number of observations, (3) and (4) present the mean value of the outcome variables for treatment and control students, column (5) indicates the difference between the means as a percentage of the control group mean and column (6) shows the p-value indicating the statistical significance of the difference across groups.

Table 6a: Descriptive Statistics across all students

Outcome Variable	Number of Observations	Treatment Students	Control Students	Difference (%)	P-Value
Self-Employment					
Home Business Ownership	1,716	50.57	31	63.13	0.000***
(%)	1659	63.9	38.1	67.72	0.000***
Any Business Ownership (%)	537	120,046	111,587	7.58	0.609
Business Income (in UGX)					
Wage-Employment					
Paid Employment (%)	1757	16.1	15.8	1.9	0.891
Wages (in UGX)	253	91,224	79,203	15.18	0.294
Income					
Monthly income across all students (in UGX)	1757	60,135	36,418	65.12	0.000***
Leadership					
Community Project Ownership	1717	36.6	16.4	123.17	0.000***
(%)	1757	77.1	74.7	3.21	0.243
School Leadership Role (%)					
Practical Skills (Min-Max)					
Business Planning (0-4)	1644	2.7	2.8	3.57	0.349
Financial Literacy (0-9)	1642	3.4	3.3	3.03	0.681

Savings (0-3)	1665	1.9	1.6	18.75	0.000***
Soft Skills					
Creativity (0-20)	1720	10.6	10.3	2.91	0.056*
Self-Efficacy (0-60)	1727	39.7	38.2	3.93	0.003***

Significance at 90% indicated by (*); at 95% by (**) and at 99% by (***)

Table 6b: Descriptive Statistics across female students

Outcome Variable	**Number of Observations**	**Treatment Students**	**Control Students**	**Difference (%)**	**P-Value**
Self-Employment					
Home Business Ownership	781	45.5	18.05	152.08	0.000***
(%)	780	63.1	28.8	119.1	0.000***
Any Business Ownership (%)	195	112,741	99,437	13.38	0.664
Business Income (in UGX)					
Wage-Employment					
Paid Employment (%)	780	13.1	11.8	11.02	0.633
Wages (in UGX)	85	87,844	60900	44.24	0.268
Income					
Monthly income (in UGX)	780	48,669	21,080	130.88	0.002**
Leadership					
Community Project Ownership	761	33.9	11.0	208.18	0.000***
(%)	780	77.9	71.4	9.1	0.040**
School Leadership Role (%)					
Practical Skills (Min-Max)					
Business Planning (0-4)	724	2.6	2.6	0	0.515
Financial Literacy (0-9)	723	3.2	3.2	0	0.896
Savings (0-3)	739	1.7	1.2	41.67	0.000***
Soft Skills					
Creativity (0-20)	761	10.6	9.9	7.07	0.011**
Self-Efficacy (0-60)	772	39.0	37.1	5.12	0.016**

Significance at 90% indicated by (*); at 95% by (**) and at 99% by (***)

5.2 Results from Regression Analysis

Tables 7 below present the findings of the more sophisticated multivariate OLS regressions with clustered standard errors according to the specifications presented in Section 4.2, for all students and for females only respectively.

The table displays regression results across all students in column (2) – (4) and across female students only in column (5) – (7). The first column indicates the number of observations, the second column the TOT regression coefficients of the treatment variable and the third column its p-value.

Table 7: Regression Statistics across all students and across female only students

	Across all students			**Across female students**		
Outcome Variable	**Number of Obs.**	**TOT**	**P-Value**	**Number of Obs.**	**TOT**	**P-Value**
Self-Employment						
Home Business Ownership (%)	1549	0.98	0.000***	684	1.39	0.000***
Business Income (in UGX)	260	41,190	0.158	72	82,505	0.063*
Wage-Employment						
Paid Employment (%)	1611	0.090	0.648	710	0.204	0.366
Wages (in UGX)	100	-3680	0.748	24	952	0.976
Income						
Income across all students (in UGX)	832	36072	0.004***	268	55,776	0.013**
Leadership						
Community Project Ownership (%)	1569	1.19	0.000***	692	1.43	0.000***
School Leadership Role (%)	1608	0.122	0.425	909	0.059	0.785
Practical Skills (Min-Max)						
Business Planning (0-4)	1508	-0.015	0.700	665	-0.049	0.389
Financial Literacy (0-9)	1508	0.088	0.427	661	0.025	0.879
Savings (0-3)	1537	0.384	0.001***	675	0.575	0.000***
Soft Skills						
Creativity (0-20)	1608	0.381	0.133	713	0.617	0.077*
Self-Efficacy (0-60)	1581	1.75	0.027**	704	2.02	0.090*

Significance at 90% indicated by (*); at 95% by (**) and at 99% by (***)

Section 6- Discussion of Results by Goal

6.1 Increased Levels of Self-Employment and Wage Employment and, subsequently, Income

For two of the five outcome indicators for this goal the target was achieved of a positive statistically significant effect experienced by treatment students compared to control students. Those outcomes which were affected as per *Educate*!'s target were the prevalence of business ownership as well as the income level across all students.

The *Educate*! program has a substantial positive impact on business ownership among the students. The effect has been found to be statistically significant at the 99% level. Treatment students have a probability of owning a business of almost 100% relative to the control students, while female treatment students have a 130% probability of business ownership compared to female control students. It becomes obvious that the participation of students in the Educate! program successfully motivates students to start a business. However, business income as well as jobs created through the business (not listed here) were not found to have been significantly affected through the program. One has to consider, though, the severe challenges in measuring business income of small business owners as indicated in section 2.2.5 as well as mentioned by McKenzie & Woodruff (2013, p. 11) implying that business income suffered of a very low sample size, rendering it difficult to detect any impact on this outcome.

Though the program seems to effectively promote business start-ups, it is disputable if the program affects the quality of the started business, measured by such proxies as business income or business employees. These results fit well with the broader literature on entrepreneurship education, arguing that it is possible to equip students with the skills required for business start-up but possibly not for business performance, and cohere with the findings of Klinger and Schündeln (2007) and Karlan and Valdivia (2011).

While the Educate! Program has an impact on self-employment, it does not show an impact on wage employment. The results indicate that the program did not promote graduates' chances of finding a salaried job nor increase their wage earnings, at least while still in school[1]. It might be the case that the program is effective in increasing self-employment but not in facilitating access to wage jobs. Yet, the lack of impact on wage employment could also be because participants were still in school

[1] Similarly to business income, the sample size for wage income was very low, potentially causing the lack of impact on derived for this variable.

at the time of the end of program survey and wage employment is not easy to attain without a greater time commitment. A follow-on survey could provide important information on this question.

It is most likely due to this strong employment effect among business owners that the overall income levels among graduates increase. In fact, graduates from the Educate! Experience program earn 36,000 UGX more per month (≈15 US$) compared to the control group. Considering the average income of control students – that is 34,000 UGX per month - graduates from the Educate! Program earn more than 100% more income in an average month. For female students the income increase is even more substantial, showing that female treatment students earn 55,000UGX more per month (≈20 US$) relative to female control students. Hence, female Educate! students earn more than 250% more money per month than the control group. Though it is to be expected that secondary school students increase their earnings after graduating from school, this impact evaluation reveals a statistical significant and substantial differential increase in earnings between treatment and control students.

These findings seem to strengthen the position of entrepreneurship education as a perceived driver of youth employment and economic growth. However, caution is recommended when considering that most entrepreneurs in Uganda are assumed to be necessity entrepreneurs. Hence, one may expect the majority of control students' businesses to fall into this category and, therefore, with *Educate*! students' business incomes not being found to differ from control students, one could conjecture that *Educate*! is primarily producing necessity entrepreneurs. These are factors to keep in mind for a follow-on analysis, since a final statement on these considerations could only be drawn from a long-term assessment of these businesses, particularly assessing their rate of business survival and increase in business income.

6.2 Improve Community Participation and Leadership Demonstration

Among the measures indicating improved community participation and leadership demonstration, a statistical significance impact has been found on ownership of community projects among the graduates. However, for holding a school leadership role no statistically significant effect was found.

With regard to the running of community projects, a positive impact was identified at the 99% level, with the regression model showing that students in the treatment group had a 120% higher probability of running a community project compared to control group students. Although there are no previous studies on the effect of entrepreneurship education on community participation for comparison, and more information is required on the activity and performance of the community projects started by the *Educate*! students, the finding for the community project ownership outcome

provides support to the assertion that leadership and entrepreneurship training has the potential to directly benefit society and, subsequently, indirectly influence the economy and employment rate. With regard to the issue of the prevalence of necessity entrepreneurs in both developing country economies as a whole and also potentially amongst *Educate*! students, this finding shows that, regardless of whether the economic activity of entrepreneurs is beneficial to the wider economy, entrepreneur development has the potential to prove beneficial above the level of the individual through other channels such as the formation of community projects.

As for the lack of impact on uptake of school leadership roles, one potential explanation might be that *Educate!* has selected students for its program that already indicate a strong leadership behaviour – an assumption supported by the relative high share of students indicating to hold leadership roles in the baseline survey - and therefore, might be leader enough for the program not to have an effect on this indicator analysed here. Furthermore, the possible larger emphasis of the program placed upon business formation and community participation, to which the findings allude, could also be used to explain the lack of impact on holding a leadership role at school, with students being led to prioritise these other endeavours. If this is the case then the negative implications of the school leadership role finding for the program's ability to improve leadership capability would be curtailed somewhat, especially when one considers that business and community project ownership could be reliably assumed to require a high degree of leadership capability.

6.3 Increased Levels of Practical and Soft Skills

Of the five composite index variables used to assess the achievement of this goal, a significantly positive impact was found upon savings behaviour and self-efficacy. No impact could be identified on the remaining practical skills, including financial literacy and business planning, and neither on the remaining soft skills, such as creativity.

The program evaluation identifies a positive and highly statistical significant effect on self-efficacy as well as savings behaviour, with significance levels at 95% and 99%, respectively. It has been established earlier that the confidence that one can act effectively to bring about desired results, more precisely the confidence in one's ability to start a new business, is possibly positively affecting business start-up. Similarly, a disciplined savings behaviour might essentially equip students with the necessary resources to start implementing their project idea. Though it may also be the case that the increase in overall income is the driving force behind the positive savings behaviour, it is less likely considering the baseline variable of overall income being found not to be statistically significant in the regression model for savings behaviour. Considering these positive findings on self-efficacy and

savings behaviour, one could conjecture that the positive impact on business and community project start-ups may have been aided by the increase in self-efficacy and savings.

In attempting to explain the lack of impact on other outcomes, such as for the business planning and financial literacy indicators, one could perhaps argue that *Educate*! may be offering no more than is already being provided by the formal education system through subjects such as Entrepreneurship Studies, which treatment and control students have access to in equal measure. However as soft skill development is supposed to provide the unique element of entrepreneurship development, this explanation cannot be offered for the lack of impact seen here on creativity. Since one might expect creativity and practical skills to positively affect business performance, the lack of impact thereof may have contributed to the lack of impact on business performance, measured by indicators including business income or business employees. This latter possibility has slight support by the baseline financial literacy variable being found to be significant at the 95% level for the business income regression models, although there is a valid concern of a possible type II error for these measures.

For Educate!, the findings for this goal are salient. A potential defence of the lack of impact seen on business performance would be that, at this stage of students' development, merely running a business or a community project is a valuable enough outcome as it facilitates learning-by-doing. Additionally, inherent challenges exist in accurately measuring practical or soft skills.

Conclusion

With the focal organisation of this paper, *Educate!*, increasing its reach annually, this paper represents a well-advised effort to reflect upon the effectiveness of its program. In a bid to do this, it has implemented a Clustered Randomised Control Trial to assess its program's impact, and the author has attempted to match the rigour of the trial design with a thorough analysis of the data it has produced, with the ultimate goal of assessing the extent to which the Educate! program has achieved the goals set for this stage of the trial, immediately after program completion.

Through utilising an OLS regression model with clustered standard errors, this analysis has found, from its most sophisticated estimates, that the program has achieved the desired target for this stage, of any statistically significant increase, in five of their twelve outcome indicators: business ownership, overall income, community project ownership, savings behaviour and self-efficacy. When female students are analysed, the results are similar, except that there are additional positive impacts on business income and creativity.

A number of immediately relevant insights can be taken by *Educate!* from this report's analysis. One main message is that their program's approach to increasing business and community project ownership is proving to be very effective, something which is also providing a boost to students' income levels. However, although we do see an impact on savings behaviour and self-efficacy, as well as creativity for girls, another lessons is that there does not seem to be evidence of the desired effect of producing students with all of the improved practical and creativity skills desired directly after program completion, something which may lead *Educate!* to reconsider some of its beliefs on how best to develop entrepreneurial ability.

Despite the useful insights this paper offers, it is clear from the amount of conjecture offered to explain its findings in the discussion section of this report that much richer information is required in order to uncover the causal chains that have produced the findings. One of the main criticisms of RCTs and quantitative impact evaluation in general is their narrow focus on attribution and subsequent neglect of opening the 'black box' of causal mechanisms (White, 2013). In an attempt to investigate this 'black box' an in-depth analysis has been conducted investigating the potential impact of several program monitoring statistics, among others student attendance, to expose the effect of certain program components on the student-level outcomes. Due to low sample size, resulting from restricting the sample to the treatment group only, the results are inconclusive. It cannot clearly be referenced which program component led to the observed impact on the student-level outcomes[2].

Alternatively, this issue could be addressed through a parallel or ex-post, complimentary collection and analysis of qualitative data (ibid). In the case of this analysis, individual or focus group interviews with *Educate*! students could provide illumination on such findings as the potential lack of impact on business performance,, contributing to the discussion around the prevalence of necessity entrepreneurs. Finally, greater insight into the types of community projects and their potential impact could be further investigated through this channel. More generally, the richer data produced could help to identify specific areas of program design and delivery that could be changed in order to improve outcomes.

In terms of future quantitative analysis of this program, the recommendation of this paper is to improve upon its weaknesses. Regarding a follow-on study, the main issue to address would be the neutrality of the analysis. Having an analysis carried out by an in-house team member - although the analyst endeavoured to remain as neutral as possible – naturally leaves the study open to accusations

[2] Results of this analysis are available upon request.

of bias. Furthermore, externally designed metrics may be more preferable than the internally developed ones used here, understanding however the inherent lack of them in the field. Finally, stronger insight would be gained on the impact on business income and job wages by increasing the sample size of any future studies.

References

Acs, Z., 2006. How is Entrepreneurship Good for Economic Growth? *Innovations,* Vol 1(1), pp. 97-107

Altman, D.G., 1985. Comparability of Randomised Groups. *The Statistician.* Vol 34, pp. 125-136

Awogbenle, A.C. & Iwuamadi, K.C., 2010. Youth Unemployment: Entrepreneurship Development Programme as an Intervention Mechanism. *African Journal of Business Management,* Vol 4(6), pp. 831-835

Azim & Kahtani, 2014. Entrepreneur Education and Training: A Summary of Literature. *Life Source Journal,* Vol 11(1), pp. 127-135

Bandiera, B., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I. & Sulaiman, M., 2012. *Empowering Adolescent Girls: Evidence from an RCT in Uganda.* Unpublished Working Paper

Barahona, C., 2010. Randomised Control Trials for the Impact Evaluation of Development Initiatives: A Statistician's Point of View. *Institutional Learning and Change Working Paper 13.*

Bhargava, S., 2007. Towards Entrepreneurship Development in the 21st Century. In S. Bhargava (Eds.) *Developmental Aspects of Entrepreneurship.* London: Sage Publications Ltd

Blanchflower, D.G. & Oswald, A.J., 1998. Entrepreneurship and the Youth Labour Market Problem: A Report to OECD. *Report to OECD, Paris.*

Boussouara, M. & Deakins, D., 1998. Learning, Entrepreneurship and the High Technology Small Firm. *Proceedings of the Enterprise and Learning Conference, Aberdeen.*

Bygrave, W.D., 1994. *The Portable MBA in Entrepreneurship.* Chichester: John Wiley

Cameron, A.C. & Miller, D.L., 2013. A Practitioner's Guide to Cluster Robust Inference. *Forthcoming in Journal of Human Resources.*

Cunningham, J.B. & Lischeron, J., 1991. Defining Entrepreneurship. *Journal of Small Business Management,* Vol 29(1), pp. 45-61

Davies, L.G. & Gibb, A.A., 1991. Recent Research in Entrepreneurship. *Proceedings of the 3rd International EIASM Workshop, Gower, London.*

DeJaeghere, J. & Baxter, A., 2014. Entrepreneur Education for Youth in Sub Saharan Africa: A Capabilities Approach as an Alternative Framework to Neoliberalism's Individualising Risks. *Progress in Development Studies,* Vol 14(1), pp. 61-76

Duflo, E., Glennerster, R. & Kremer, M., 2007. Using Randomization in Development Economics Research. *Handbook of Development Economics,* Vol 4, pp. 3895-3962.

Dunning, T. & Hyde, S., 2008. The Comparing of Experimental Data: Comparing Techniques. *Annual Meetings of the American Political Science Association, Boston, MA*

Duval-Couetil, N., 2013. Assessing the Impact of Entrepreneurship Education Programs: Challenges and Approaches. *Journal of Small Business Management,* Vol 51(3), pp. 394-409

Educate!, 2013. *Annual Performance Report* [Online] Available at: http://www.experienceeducate.org/annual-reports/ (Accessed on 5th November 2014)

Ehrlich, S., De Noble, A. & Jung, D., 2000. Entrepreneurial Self-Efficacy: The Development of a Measure and its Relationship to Entrepreneurial Action. *Frontiers for Entrepreneurship Reseach: University of San Diego*

Fairlie, R.W., Karlan, D. & Zinman, J., 2012. Behind the Gate Experiment: Evidence on Effects of and Rationales for Subsidised Entrepreneurship Training. *No. w17804. National Bureau of Economic Research*

Farsad, H., 2002. Integrated Entship Education in Botswana, Uganda and Kenya. *Report for National Institute of Technology, Oslo, Norway*

Fayolle, A., 2008. Entrepreneurship Education at a Crossroads: Towards a More Mature Teaching Field. *Journal of Enterprising Culture,* Vol. 16(4), pp. 325-337

Fiet, J.O., 2000. The Theoretical Side of Teaching Entrepreneurship Theory. *Journal of Business Venturing*, Vol. 16(1), pp. 1-24.

Glaub, M. & Frese, M., 2011. A Critical Review of the Effects of Entrepreneurship Training in Developing Countries. *Enterprise Development and Microfinance,* Vol 22(4), pp. 335-353

Goldman, M., 2008. *Statistics for Bioinformatics* [Online] Available at: http://www.stat.berkeley.edu/~mgoldman/Section0402.pdf (Accessed on 21st June 2014)

Gorman, G., Hanlon, D. & King, W., 1997. Some Research Perspectives on Entrepreneur Education, Enterprise Education and Education for Small Business Management: A Ten-Year Literature Review. *International Small Business Journal,* Vol 15(3), pp. 56-77

Hansemark, O., 1998. The Effects of an Entrepreneurship Programme on Need for Achievement and Locus of Control of Reinforcement. *International Journal of Entrepreneurial Behaviour & Research,* Vol 4(1), pp. 28-50

Harper, M. & Finnegan, G., 1998. *Value for Money? Impact of Small Enterprise Development.* London: Intermediate Technology Publications

Heckman, J.J., Lochner, L. & Taber, C., 1998. General Equilibrium Treatment Effects: A Study of Tuition Policy. *The American Economic Review,* Vol. 88(2), pp. 381-386

Henry, C., Hill, F. & Leitch, C., 2005. Entrepreneur Education and Training: Can Entrepreneurship Be Taught? Part 1. *Education + Training,* Vol 47(2), pp. 98-111

Hutchison, D. & Styles, B., 2010. *A Guide to Running Randomised Controlled Trials for Educational Researchers.* Slough: NFER

Imbens, G.M. & Wooldridge, J.M., 2008. Recent Developments in the Econometrics of Program Evaluation. *No. w14251. National Bureau of Economic Research.*

International Youth Foundation, 2013. Testing What Works in Youth Employment: Evaluating Kenya's Ninaweza Program. *Volume 1: A Summative Report*

Isaakidis,P. & Ioannidis, J., 2003. Evaluation of Cluster Randomised Controlled Trials in Sub-Saharan Africa. *American Journal of Epidemiology,* Vol 158(9), pp. 921-926.

Jack, S.L. & Anderson, A.R., 1998. Entrepreneurship Education Within the Condition of Entreprenology, *Proceedings of the Conference on Enterprise and Learning, Aberdeen.*

Jacob, R. & Zhu, P., 2009. New Empirical Evidence for the Design of Group Randomized Trials in Education. *MDRC Working Paper.*

James-Wilson, D., 2008. Youth Livelihoods Program Development Guide. *USAID/EGAT/Office of Education: Washington DC.*

Jamieson, I., 1984. Education for Enterprise, in Watts, A.G. and Moran, P. (Eds) *Education for Enterprise*, CRAC/Ballinger, Cambridge, pp. 19-27.

Karlan, D. & Valdivia, M., 2011. Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions. *Review of Economics and Statistics,* Vol 93(2), pp. 510-527

Kennedy, J. & Peterman, N.E., 2003. Enterprise Education: Influencing Students' Perceptions of Entrepreneurship . *Entrepreneurship*: *Theory & Practice*, Vol 28, pp. 129-145.

Khandker, S.R., Koolwal, G.B. & Samad, H.A., 2010. *Handbook on Impact Evaluation: Quantitative Methods and Practices.* World Bank Publications.

King, E & Hill, M.A., ,1997, eds. *Women's Education in Countries: Barriers, Benefits and Policies.* World Bank Publications.

Kling, J.R. & Liebman, J.B., 2004. Experimental Analysis of Neighbourhood Effects on Youth. *Princeton IRS Working Paper 483*

Klinger, B. & Shundeln, M., 2007. Can Entrepreneurial Activity be Taught? Quasi-Experimental Evidence from Central America. *CID Working Paper No. 153*

Kolvereid, L. & Moen, O., 1997. Entrepreneurship Among Business Graduates: Does a Major in Entrepreneurship Make a Difference? *Journal of European Industrial Training,* Vol 21(4), pp. 154-160

Lewis, S.C. & Warlow, C.P., 2004. How to Spot Bias and Other Potential Problems in Randomised Controlled Trials. *J Neurol Neurosurg Psychiatry,* Vol 75, pp. 181-187

Mahlberg, T., 1996. Evaluating Secondary School and College Level Entrepreneurial Education-Pilot Testing a Questionnaire. *The Internationalising Entrepreneurship Education and Training Conference. The Netherlands: Arnhem/University of Nijmegen.*

Edyegu, D., 2012. Uganda's Museveni Admits Youth Unemployment is Out of Control. *Mail and Guardian* [Online] Available at: http://mg.co.za/article/2012-10-10-ugandas-museveni-admits-youth-unemployment-is-out-of-control (Accessed on 2nd August 2014)

Mair, J. & Marti, I., 2006. Social Entrepreneurship Research: A Source of Explanation, Prediction and Delight. *Journal of World Business.* Vol 41(1), pp. 36-44

Martinez, A.C., Levie, J., Kelley, D.J., Saemundsson, R.J., Schott, T., 2010. A Global Perspective on Entrepreneurship Education and Training. *Global Entrepreneurship Monitor: Special Report.*

Mathisen, T., Johansen, V., Mathisen, S., 2011. *Evaluation of Entrepreneurship in Education: A Program Between Junior Achievement Uganda and Junior Achievement Norway*. Lillehammer: Eastern Norway Research Institute

Matlay, H., 2005. Researching Entrepreneurship and Education: Part 1: What is Entrepreneurship and Does it Matter? *Education + Training,* Vol 47(8/9), pp. 665-677

McKenzie, D. & Woodruff, C., 2013. What Are We Learning from Business Training and Entrepreneurship Evaluations Around the Developing World? *The World Bank Research Observer,* Vol 29(1), pp. 48-82

Michael, S.C., 2007. Entrepreneurship, Growth and Adam Smith. *Strategic Entrepreneurship Journal,* Vol 1, pp. 287-289

Miller, A., 1987. New Ventures: A Fresh Emphasis on Entrepreneurial Education. *Survey of Business*, Vol. 23(1), pp. 4-9.

Mwasalwiba, E., 2010. Entrepreneur Education: A Review of its Objectives, Teaching Methods and Impact Indicators. *Education + Training,* Vol 52(1) pp. 20-47

Nafukho, F. & Muyia, M., 2010. Entrepreneurship and Socioeconomic Development in Africa: A Reality or Myth? *Journal of European Industrial Training,* Vol 34(2), pp. 96-109

Nelson, R. & Johnson, S., 1997. Entrepreneurship Education as a Strategic Approach to Economic Growth in Kenya. *Journal of Industrial Teacher Education,* Vol 35, pp7-21

O'Halloran, S., 1996. *Lecture 5: Model Checking*. [Online] Available at: http://www.columbia.edu/~so33/SusDev/Lecture_5.pdf (Accessed on 3rd October 2014)

Oosterbeek, H., van Praag, M. & Ijsselstein, A., 2010. The Impact of Entrepreneurship Education on Entrepreneurship Skills and Motivation. *European Economic Review,* Vol 54, pp. 442-454

Orford, J., Herrington, M. & Wood, E., 2004. South Africa Report. *Global Entrepreneurship Monitor.*

Oyugi, J., 2014. Effectiveness of the Methods of Teaching Entrepreneurship Courses to Developing Self-Efficacy and Intention among University Students in Uganda. *International Journal of Social Sciences and Entrepreneurship,* Vol 1(11) pp. 1-22

Pittaway, L. & Cope, J., 2006. Entrepreneur Education: A Systematic Review of the Evidence. *International Small Business Journal,* Vol 25(5), pp. 479-510

Pocock, S.J., Assmann, S.E., Enos, L.E. & Kasten, L.E., 2002. Subgroup Analysis, Covariate Analysis and Baseline Comparisons in Clinical Trial Reporting: Current Practice and Problems. *Statistics in Medicine,* Vol 21(19), pp. 2917-2930.

Puffer, S., Torgerson, D. & Watson, J., 2003. Evidence for Risk of Bias in Cluster Randomised Trials: Review of Recent Trials Published in Three General Medical Journals. *BMJ,* Vol 327(7418), pp. 785-789

Premand, P., Brodmann, S., Almeida, R., Grun, R. & Barouni, M., 2012. Entrepreneurship Training and Self-Employment Among University Graduates: Evidence from a Randomised Trial in Tunisia. *IZA DP No. 7079*

Pretorius, M.,2008. Assessment of Entrepreneur Education: A Pilot Study. *South African Journal of Entrepreneurship and Small Business Management.* Vol. 1(1), pp. 1-20

Primo, D.M., Jacobsmeier, M.L & Milyo, J., 2007. Estimating the Impact of State Policies and institutions with Mixed-Level Data. *State Politics & Policy Quarterly,* Vol 7(4), pp. 446-459

Sankoh, A.J., Huque, M.F. & Dubey, S.D., 1997. Some Comments on Frequently Used Multiple Endpoint Adjustment Methods in Clinical Trials. *Statistics in Medicine.* Vol 16(22), pp. 2529-2542

Seelos, C. & Mair, J., 2005. Social Entrepreneurship: Creating New Business Models to Serve the Poor. *Business Horizons,* Vol 48(3), pp. 241-246

Shaffer, J.P., 1995. Multiple Hypothesis Testing. *Annual Review of Psychology.* Vol 46, pp. 561-584

Shepherd, D.A. & Douglas, E.J., 1996. Is Management Education Developing or Killing the Entrepreneurial Spirit? *Proceedings of the Internationalising Entrepreneurship Education and Training Conference, Arnhem*

The State of Uganda, 2013. The State of Uganda Population Report: Population and Social Transformation: Addressing the Needs of Special Interest Groups.

Steinacher, R & Collazos, A., 2013. Baseline and Program Implementation Report: Educate! Experience Evaluation. *Internal Report Produced by Educate! Staff*

Subba Rao, S.R. & Durga Prasad, Ch.S., 2007. Entrepreneurship: An Effective Means to Promote Employment. In S. Bhargava (Eds.) *Developmental Aspects of Entrepreneurship.* London: Sage Publications Ltd

Timmons, J.A., Muzyka, D.F., Stevenson, H.M. & Bygrave, W.D., 1987. Opportunity Recognition: The Core of Entrepreneurship, in Churchill, N. (Ed.), *Frontiers of Entrepreneurial Research, Babson College*. Babson Park: Massachusetts, pp. 42-49.

Torgerson, C. & Torgerson, D., 2001. The Need for Randomised Control Trials in Educational Research. *British Journal of Educational Studies,* Vol 49(3), pp. 316-328.

Wears, R.L., 2002. Advanced Statistics: Statistical Methods for Analysing Cluster and Cluster-Randomised Data. *Academic Emergency Medicine,* Vol 9(4), pp. 330-341.

White, H., 1980. A Heteroskedasticity-Consistent Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica,* Vol 48, pp. 817-838

White, H., 2013. The Use of Mixed Methods In Randomised Control Trials. In D.M Mertens & S.Hesse-Biber (Eds.), *Mixed Methods and Credibility of Evidence in Evaluation. New Directions for Evaluation,* Vol 138, pp. 61-73

Wiegratz, J., 2009. *Uganda's Human Resource Challenge: Training, Business Culture and Economic Development.* Kampala: Fountain Publishers

Wooldridge, J.M., 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge: The MIT Press.

Wooldridge, J.M., 2013. *Introductory Econometrics: A Modern Approach.* UK: Cengage Learning.

World Bank, 1993. *Uganda: Growing Out of Poverty*. Washington DC: World Bank

World Bank, 2014. *Unemployment, Youth Total (% of Total Labor Force Aged 15-24) (National Estimate).* [Online] Available at: http://data.worldbank.org/indicator/SL.UEM.1524.NE.ZS (Accessed on 6th August 2014)

World Economic Forum, 2009. Educating the Next Wave of Entrepreneurs: Unlocking Entrepreneurial Capabilities to Meet the Global Challenges of the 21st Century. *A Report of the Global Education Initiative.*

Young, J.E., 1997. Entrepreneurship Education and Learning for University Students and Practicing Entrepreneurs. In D.L. Sexton & R.W. Simlor (Eds.). *Entrepreneurship 2000*. Chicago, IL: Upstart Publishing.

The Youth Employment Network, 2013. Getting it Right from the Baseline. *YEN Briefs Issue*

The following Appendix along with the full data set and analysis is also available upon request.

Appendix 1: List of Districts and Schools Included in Trial
Appendix 2: Power Calculations
Appendix 3: End of Program Survey
Appendix 4: Construction of Composite Metric Scores
Appendix 5: Educate! Data Analysis Guidelines on Missing Values