Pre-Analysis Plan for "The Impacts of Female Education: Evidence from Malawian Secondary Schools"

Authors: Syngjoo Choi, Booyuel Kim, Hyuncheol Bryant Kim and Cristian Pop-Eleches.

May, 2016

Section 1: Introduction

This pre-analysis plan describes the hypothesis and specifications that will be used to understand the impact of female education in the context of secondary schools in Malawi. This project aims to produce two papers: the first paper will look at impacts of an education intervention on rationality, risk preferences, and time preferences. The second paper will study the impact of the education intervention on risky sexual behaviors and sexually transmitted diseases (HIV and HSV2). This plan was written while the follow-up data was collected and before any data was analyzed. The organization of the plan is as follows: In section 2 we describe the study, the sample selection and the data that was and will be collected for this project. In Sections 3 and 4 we describe the hypotheses to be tested for this study and the specifications that will be performed during data analysis. In the appendix we provide a number of additional information about the study and data.

Section 2: Overview of the study

2.1 Motivation

Education is considered one of the most important determinants of human capital development and it is widely agreed that improved education is correlated with increased wages later in life at the micro-level (Blau and Kahn 2005) and may lead to more economic growth of a country at the macro-level (Hanushek and Woessmann 2012). However, there is a lot less evidence on the specific underlying mechanisms to connect educational interventions to test scores and later adulthood outcomes.

The first purpose of this project is to understand the impact of an education intervention in secondary schools in Malawi on rationality, risk preferences, time preferences. The second purpose of the project is to understand the impact of the same education intervention on risky sexual behaviors and sexually transmitted diseases (HIV and HSV2). This project evaluates the Girls Education Support Program (GESP) for 9th – 11th graders implemented by the Africa Future Foundation in four catchment districts of rural Lilongwe. These four districts are Chimutu, Chitukula, Tsabango, and Kalumba. For the school year 2012-3, all girls in the treatment classroom received tone-year tuition support and monthly cash stipends. School tuition and fees per semester on average were 3,400 Malawi Kwacha and were directly deposited to each school's accounts, and monthly payments of 300 Kwacha were distributed to treated students. The overall amount of support was around $70/year. Given the limited resources of the Africa Future Foundation, that could not cover scholarships for all girls, the evaluation used randomly selected classrooms within grades from a sample of schools and compared the outcomes with girls that were not selected to receive scholarships.

2.2. Sample Selection

The evaluation sample is based on 3,997 female students. About half the female students (2,102) were selected into the girls' education support program (treatment group), and the others (1,895) were in the control group. The original sample of students comes from a 124 classrooms in 33 public secondary schools in Malawi. For the current study we have selected the 2,811 students who were in grade 9 and 10 in 2011 (since those in grade 11 are harder to track). These students come from 83 classrooms grade-school in 33 schools.

2.2.1 Catchment district selection

We selected the four catchment districts (Chimutu, Chitukula, Tsabango, and Kalumba) because they are the catchment area of the Daeyang Luke Hospital in Lilongwe, Malawi, which was the partner hospital of the Africa Future Foundation for this project.

2.2.2 School Selection

We invited all the 33 public schools in these districts to participate in the girls' education programs, but we excluded the private boarding schools located in the catchment area.

2.2.3 Assignment to Treatment

Based on a survey of all 33 schools, we identified 124 classrooms. Assignment of the treatment to classrooms was done randomly in the project office by a computer random number generator.

2.3 Sources of data

The primary data sources are a baseline survey conducted in October, 2011 to May, 2012, a short term follow-up survey implemented in January to June 2013, and the main long term follow-up survey that has started in October 2015 and is scheduled to be finished in August, 2016.

2.3.1 Baseline survey

The baseline survey was implemented between October, 2011 and May, 2012. 7971 secondary students (3,997 girls and 3,974 boys) participated to our baseline survey. 74.4% of total 10,715 students in the school roll-call lists completed the baseline survey. In Appendix Table 1, we show that the randomization was generally successful since there are very few statistically significant differences between the treatment and control groups, for both the full sample as well as the target sample for this evaluation (girls in form 9 and 10). The baseline survey is separately attached.

2.3.1 Short-term follow-up survey

The short-term follow-up survey will not be used in the analysis, except for understanding the impact of the intervention on a range of educational outcomes in order to understand the education impact of the Girls Education Support Program (GESP). The follow-up survey is separately attached.

2.3.2 Long-term follow-up survey

The long-term follow-up survey started in October of 2015 and will be completed by August of 2016. The long-term follow-up survey is separately attached.

Section 3: Hypothesis

Our long-term follow-up survey has collected a large number of outcome variables in order to understand how the experimentally induced changes in educational achievement affect 1) measures of rationality and preferences and 2) HIV and HSV2 infection.

Section 3.1: Impacts of Education on Rationality and Preference

The Girls Education Support Program (GESP) may have an average treatment effect on rationality, risk preferences and time preferences. Below we describe the hypothesis to be tested:

Hypothesis 3.1.1: GESP enhances average level of economic rationality. We use revealed preference analysis to check whether subjects' behavior comply with utility maximization hypothesis in each domain of decisions under risk and inter-temporal choices. This will be done by checking whether individual behavior is consistent with GARP (Generalized Axiom of Revealed Preferences), and measuring the extent to which it violates GARP.

We use indices of measuring GARP violations in each domain:

1.) Afriat's Critical Cost Efficiency Index (CCEI) measures the amount by which each budget constraint must be adjusted in order to remove all violations of GARP It ranges between zero and one. The closer the CCEI is to one, the smaller the adjustment of the budget constraints required to remove all violations and thus the closer the data are to satisfying GARP.

2.) Varian (1991) measure refines Afriat's CCEI by providing a measure that reflects the minimum adjustment required to eliminate the violations of GARP associated with each price-choice observation.

This will be separately done in domains of decision making under risk as well as of inter-temporal choices.

Hypothesis 3.1.2: GESP has an impact on risk preferences. For risk preferences, we do not have a prior belief on the direction of the change. The following variables are part of this domain: experimental data of choices under risk will allow us to measure two distinct parameters of risk attitudes-utility curvature and probability weighting. This will be based on parametric estimation of rank-dependent utility model (Quiggin, 1981; Schmeidler, 1989). We will also use a single, nonparametric measure of risk attitudes.

Our experimental approach of measuring risk preferences will allow us to have the following three measures of risk preferences.

1) Average demand share on a cheaper asset: Risk neutrality or risk seeking behavior will predict this measure to be 1. The more risk averse the individual is, the closer this measure is to 1/2.

2) A parametric measure of probability weighting

3) A parametric measure of utility curvature

We rely on a structural estimation of rank-dependent utility model to estimate probability weighting and utility curvature.

Hypothesis 3.1.3: GESP leads female students to be more patient. Using experimental data of intertemporal choices, we will measure two parameters of time impatience-present bias and standard discount rate. This will be based on parametric estimation of quasi-hyperbolic discounting utility model (Laibson, 1997). Also, a single, nonparametric measure of time impatience will be used.

Specifically, we consider the following three measures:

1) Average demand share on a sooner payment date: The more impatient the individual is, the higher this measure.
2) A parametric measure of present bias
3) A parametric measure of exponential time discount rate

We rely on a structural estimation of quasi-hyperbolic discounting utility model to estimate present bias and standard discount rate.

3.2 Causal chain of process and mechanisms

Hypothesis 3.2.1: The impacts of GESP will be greater on outcomes if the educational intervention had a larger impact on educational attainment.

Our main measure of educational attainment will be the number of completed education years (measured at the long-term follow-up survey).

Section 3.2: Long-term Impacts of Education on Sexually Transmitted Diseases infection measured by HIV and HSV2 infection

Hypothesis 3.2.1: GESP may affect HIV and HSV2 infection in the long run.

The long-term Impacts of cash transfer on schooling for Sexually Transmitted Diseases (STDs) infection can be theoretically ambiguous. In the short-run, schooling may prevent STD infection; spending time during the day at school itself protects students from enagaging in sex activities. School-based HIV prevention programs can help to promote safe sexual behaviors and reduce stigma on HIV (Juke et al 2008), and general education may also encourage students to protect themselves from HIV infection. However, cash transfer for schooling might increase HIV and HSV infection in the long-run if delayed marriage due to staying in schools allows students to have more life time sexual partners.

In fact, Sarah et al (2010 and 2011) found that cash transfer for schooling could delay sexual debut, pregnancy, and marriage and prevent HIV and HSV2 infection in the short-run (18 months). However the results could be spurious since incidences of marriage, pregnancy, and HIV and HSV2 in the follow-up survey were very low, given that it only measured the short-term effects.

Our outcome variables are HIV infection and HSV2 infection (measured by serum IgG). We expect that we have more statistical power to detect difference in HSV2 infection than HIV infection since prevalence of HSV2 is higher than that of HIV. We may be under powered to detect statistical differences in the HIV prevalence rate given that the HIV infection rate is relatively low. According to Malawi AIDS Response Progress Report 2015, percentage of women aged 15-24 who are living with HIV is 3.6%. As a result, the main outcome variable for the study will be the HSV2 infection rate. The secondary outcome will be the HIV infection rate.

Hypothesis 3.2.2: HIV and HSV2 infection is higher among those timing of marriage is delayed.

We try to explore the potential mechanism through which schooling affects STDs infection. As mentioned, cash transfer for schooling might increase STDs infection in the long-run if it delays marriage. We, therefore, test whether HIV and HSV2 infection is higher among those for whom the timing of marriage is delayed through the GESP program. Additionally, we will explore age of the partner, sexual behaviors, attitudes toward condom use, and non-cognitive traits such as self-esteem and confidence to explain the mechanisms further.

Section 4: Estimation:

4.1. Estimation of treatment effects

Our main analysis will focus on the reduced-form estimation of GESP impacts on outcomes. We will be estimating the following simple model using ordinary least squares:

(1) $Y_{icg} = \beta_0 + \beta_1 ES_{cg} + \beta_2 X_{icg} + \delta_g + \varepsilon_{icg}$

The variable Y_{ic} is one of our outcomes discussed previously for student *i* in classroom *c* at grade *g*. The variable ES_{cg} is an indicator for whether classroom *c* at grade *g* was assigned for education support program. The coefficient β_1 captures the intent-to-treat effect of being selected to a treatment classroom among girls in our experimental sample. X_{icg} is a vector that includes the following socio-demographic controls, measured at baseline: age, orphan status, father's education, mother's education, father's job, mother's job, household assets, and school type.

All specification will also include grade fixed effects, δ_g since the randomization was implemented at the classroom level within grade. Standard errors will be clustered at the classroom level. In some specifications we will also control for the effect of two interventions (HIV/AIDS education and circumcision for boys) and their interactions, since the broader project consists of GESP for girls and male circumcision for boys.

4.2. Estimation of heterogeneous treatment effects: heterogeneous treatment effects will be estimated by an equation that interacts treatment status as well as all the control variables with the variable of interest.

4.3 Dealing with testing for multiple outcomes for standardized treatment effects and adjustments for multiple inference.

In order to account for multiple hypotheses testing, we follow the pre analysis plan provided by Finkelstein et al. (2010). A first approach used in Finkelstein et al. (2010) is to group outcome measure into domains. For our analysis of the first paper project, the three domains are rationality, time preference, and risk preference. Next, we sign the outcomes in each domain, and take a standardized treatment effect in that domain, as suggested in Kling, Liebman, and Katz (2007). In order to account for multiple inference within our four domains, we will use family-wise error rate adjusted p-values using the Westfall and Young step-down resampling method. To account for multiple hypotheses testing by analyzing the treatment effect heterogeneity, we will use the method to minimize the false non-discovery rate (Benjamini and Hochberg (1995), Fink, McConnell, and Vollmer (2010). Treatment effect heterogeneity will be limited to the overall standardized domain outcome for the four domains (rationality, time preference, risk preference, STDs infection).

4.4. Survey attrition

We plan to potentially use a more extensive tracking exercise if the budget allows after the conclusion of the standard survey as in Thomas et al. 2001 and Thomas et al. 2012. We will perform a probability-reweighting of the data and will also calculate an effective survey follow-up rate.

Specifically, we will estimate the following equation to understand attrition:

(2) $AT_{icg} = \beta_0 + \beta_1 ES_{cg} + \beta_2 X_{icg} + \delta_g + \varepsilon_{icg}$

where AT_{icg} is an attrition indicator and the other variables are the same as those defined in equation (1). Standard errors will also be clustered at the classroom level. If the treatment status variable is not significant at the 5% level, we will do our estimation as described above without making any changes. If the treatment status variable is significant at the 5% level, we will perform a bounding exercise as suggested in Lee (2009).

4.5 Missing data from item non-response

No imputation for missing data from item non-response will be performed at follow-up.

4.6 Variables with limited observations

In case survey questions have the same value for more than 95% of observations within our follow-up sample, we will omit them from the analysis.

Bibliography

Baird, Sarah, et al. "The short‐term impacts of a schooling conditional cash transfer program on the sexual behavior of young women." Health economics 19.S1 (2010): 55-68.

Baird, Sarah J., et al. "Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial." The Lancet 379.9823 (2012): 1320-1329.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 289-300.

Blau, F. D., & Kahn, L. M. (2005). Do cognitive test scores explain higher US wage inequality?. Review of Economics and statistics, 87(1), 184-193.

Coffman, K. B., Coffman, L. C., & Ericson, K. M. M. (2013). The size of the LGBT population and the magnitude of anti-gay sentiment are substantially underestimated (No. w19508). National Bureau of Economic Research.

Fink, G., McConnell, M., & Vollmer, S. (2010). Heterogeneous Treatment Effects and Multiple Hypothesis Testing in Field Experiments. Working Paper.

Finkelstein et al. (2010). The short-run impact of extending public health insurance to low income adults: evidence from the first year of The Oregon Medicaid Experiment: Analysis Plan.

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. Journal of Economic Growth, 17(4), 267-321.

Jukes, Matthew, Stephanie Simmons, and Donald Bundy. "Education and vulnerability: the role of schools in protecting young women and girls from HIV in southern Africa." AIDS 22 (2008): S41-S56.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. Econometrica, 75(1), 83-119.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. The Quarterly Journal of Economics, 443-477.

Quiggin, J. (1981). Risk perception and risk aversion among Australian farmers. Australian Journal of Agricultural Economics, 25(2), 160-69.

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. Econometrica, 57(3), 571-587.

Thomas, D., Frankenberg, E., & Smith, J. P. (2001). Lost but not forgotten: Attrition and follow-up in the Indonesia Family Life Survey. Journal of Human Resources, 556-592.

Thomas, D., Witoelar, F., Frankenberg, E., Sikoki, B., Strauss, J., Sumantri, C., & Suriastini, W. (2012). Cutting the costs of attrition: Results from the Indonesia Family Life Survey. Journal of Development Economics, 98(1), 108-123.

Varian, H. (1991). Goodness-of-Fit for Revealed Preference Tests. University of Michigan CREST Working Paper # 13.

Appendix Table

Table 1. Baseline statistics and randomization balance

Dependent Variable:	Avg. (s.d)	HIV/AIDS Education	Male Circumcision	Male Circumcision (eligible boys)	Male Circumcision (ineligible girls)	Girls Scholarship	Girls Scholarship (eligible girls)	Girls Scholarship (ineligible boys)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Full sample								
Age (year)	16.16	-0.008	0.010	-0.006	-0.007	0.000	0.003	0.001
	(1.856)	(0.006)	(0.008)	(0.007)	(0.011)	(0.008)	(0.011)	(0.008)
Orphan	0.054	-0.009	-0.016	-0.027	-0.006	-0.010	-0.037	0.013
	(0.227)	(0.025)	(0.024)	(0.037)	(0.029)	(0.024)	(0.032)	(0.035)
Father's tertiary education	0.198	0.009	0.010	-0.011	0.020	0.014	0.047**	-0.021
	(0.399)	(0.019)	(0.017)	(0.027)	(0.022)	(0.019)	(0.022)	(0.029)
Mother's tertiary education	0.082	-0.020	0.009	-0.014	0.038	0.033	-0.024	0.104***
	(0.274)	(0.027)	(0.025)	(0.038)	(0.030)	(0.024)	(0.030)	(0.037)
Father's white-collar job	0.256	-0.006	-0.006	0.027	-0.033	-0.002	-0.031	0.033
	(0.436)	(0.016)	(0.016)	(0.022)	(0.021)	(0.017)	(0.023)	(0.024)
Mother's white-collar job	0.106	0.027	0.005	-0.028	0.020	-0.031	-0.017	-0.044
	(0.307)	(0.025)	(0.024)	(0.034)	(0.029)	(0.024)	(0.027)	(0.035)
Household Assets	7.59	0.001	0.002	0.003	0.000	0.004	0.003	0.004
(0-16)	(3.455)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Conventional School	0.245	0.051	-0.066	-0.011	-0.121	0.087	0.086	0.089
	(0.430)	(0.083)	(0.084)	(0.081)	(0.095)	(0.103)	(0.102)	(0.105)
p-value of joint F-test		0.855	0.768	0.720	0.735	0.720	0.325	0.172
Observations		7,957	7,957	3,964	3,993	7,957	3,993	3,964
R-squared		0.070	0.018	0.019	0.030	0.045	0.055	0.041

Table 2. Baseline statistics and randomization balance, female 9th and 10th graders

Dependent Variable:	Avg. (s.d)	HIV/AIDS Education	Male Circumcision	Male Circumcision (eligible boys)	Male Circumcision (ineligible girls)	Girls Scholarship	Girls Scholarship (eligible girls)	Girls Scholarship (ineligible boys)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B. Grade 9 and 10 sample								
Age (year)	15.7211	-0.013	0.013	0.004	-0.013	0.001	0.011	0.000
	(1.7121)	(0.009)	(0.009)	(0.009)	(0.014)	(0.010)	(0.014)	(0.011)
Orphan	0.0474	-0.029	-0.033	-0.044	-0.028	-0.006	-0.057	0.041
	(0.2124)	(0.029)	(0.027)	(0.043)	(0.040)	(0.031)	(0.042)	(0.045)
Father's tertiary education	0.1962	0.009	0.033	0.041	0.011	-0.000	0.036	-0.038
	(0.3972)	(0.023)	(0.022)	(0.032)	(0.025)	(0.023)	(0.027)	(0.035)
Mother's tertiary education	0.0832	-0.032	-0.013	-0.056	0.046	0.054*	-0.003	0.134***
	(0.2763)	(0.034)	(0.029)	(0.047)	(0.033)	(0.029)	(0.036)	(0.045)
Father's white-collar job	0.2543	-0.008	-0.022	-0.004	-0.035	0.030	-0.004	0.068**
	(0.4355)	(0.019)	(0.018)	(0.026)	(0.025)	(0.020)	(0.027)	(0.030)
Mother's white-collar job	0.1048	0.023	0.030	-0.014	0.038	-0.044	-0.029	-0.061
	(0.3063)	(0.031)	(0.026)	(0.045)	(0.030)	(0.029)	(0.032)	(0.044)
Household Assets	7.4823	0.003	0.001	0.002	-0.000	0.001	-0.001	0.003
(0-16)	(3.4528)	(0.006)	(0.007)	(0.007)	(0.008)	(0.009)	(0.009)	(0.009)
Conventional School	0.2349	-0.040	-0.055	0.058	-0.175	0.155	0.150	0.160
	(0.424)	(0.097)	(0.099)	(0.091)	(0.114)	(0.131)	(0.130)	(0.132)
p-value of joint F-test		0.7177	0.2934	0.7584	0.1957	0.3485	0.5011	0.0555
Observations		5469	5469	2661	2808	5469	2808	2661
R-squared		0.0635	0.0170	0.0204	0.0410	0.0326	0.0351	0.0370

Notes: This sample consists of 7,971 students (in Panel A) and a subsample of 5,478 students (in Panel B) who were interviewed at the baseline survey. Parent's tertiary education is 1 when they graduate from a 2-year college or a 4-year university. Parent's white-collar job is coded as 1 when they have a professional or government job. Household Assets are defined as the total number of assets they own from 16 asset questions. Conventional school is 1 when a student is enrolled in a conventional secondary school. Columns (2) - (8) show randomization balance for three different interventions. Columns (4) and (5) show randomization balance of male circumcision intervention for eligible boys and ineligible girls while columns (7) and (8) show the balance of girls scholarship intervention for eligible girls and ineligible boys, respectively. Robust standard errors clustered by classroom are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1