Increasing students' aspirations: the impact of a role model on students' educational attainment Pre-Analysis Plan Amendment

Emma Riley *

July 18, 2017

1 Introduction

This analysis plan amendment pertains to additional data collection associated with the trial "Increasing students' aspirations: the impact of a role model on students' educational attainment."[1] In this trial, I test whether students seeing an aspirational movie, *Queen of Katwe*, increases their performance on national exams, as set out in the original pre-analysis plan.

Additionally, as outlined in this amendment, I will test whether seeing the movie *Queen of Katwe* also increases the likelihood that students obtain a government scholarship and entry to University. I was not aware of the potential to collect these additional outcomes when writing the original pre-analysis plan. Since then, when investigating what other data could be collected, these additional outcomes became apparent as being publicly available from the Ministry of Education. I also discuss the collection of mock exam results taken before the trial, which I use as baseline exam results, allowing an ANCOVA specification to be run. These mock exam results were not collected before the trail or before the lodging of the original pre-analysis plan as, again, I was not aware that all schools sat mock exams or that they would have copies of these exams that they could make available to me.

*University of Oxford. emma.riley@economics.ox.ac.uk

[1]See http://www.socialscienceregistry.org/trials/1832 for the trial registration and https://www.socialscienceregistry.org/docs/analysisplan/957/document for the original pre-analysis plan

This amendment should be read alongside the original pre-analysis plan. All analysis detailed in this amendment will be carried out upon completion of the additional data collection in August 2017.

2 Additional Outcomes of Interest

Additional data on outcomes relating to continuation to higher education have been collected during summer 2017. This data will be used to make the following additional outcome variables:

1. An indicator for whether the student obtained a government scholarship (for students in the S6 class at the start of this trial)
2. An indicator for whether the student gained entry to University (for students in the S6 class at the start of this trial)

These outcomes are being obtained from the Ugandan National Council of Higher Education who hold records on all University entry and determine scholarship awards. These records are publicly available.

3 Mock exam results

Since the original pre-analysis plan was lodged, I have been able to request from schools data on mock exam results for all the students in the study.

Mock exams were performed in all the schools during August 2016 before the schools were approached in September and the randomised trial began in October 2016. These therefore measure students' educational ability before treatment. I will use the mock exam results as a control variable in regression specifications and for heterogeneity analysis by prior exam performance, as detailed below.

At S4 level, mock exams are taken in Maths and English subjects only. The mock exams are graded in the same manner as the main exams, with a score from 1-9 with 1 being the best score and 9 the worst. As for the outcome grades, scores will be inverted so that a 9 becomes 0 and a 1 becomes 8. This is so that a higher score can be interpreted as a better performance

As for the outcome exam results, standardized mock test scores will be created for each subject by subtracting the mean and dividing by the standard deviation of the control group. An overall aggregate of exam performance will be calculated by summing standardised test scores across the Maths and English scores and renormalising.

This gives three mock variables at S4 level:

1. The **standardised Maths** mock score
2. The **standardised English** mock score
3. The **standardised aggregate** of the English and Maths scores

At S6, mocks are taken in the 3 chosen principal subjects as well as the subsidiary and general paper. The subsidiary subjects and general paper are graded on a 1-9 scale, with 1 being the best and 9 the worst grade. Grades 7 and above are fails. Any student achieving a 6 or below on a subsidiary paper or the general paper gets one point. The principal papers are marked on a A,B,C scale, with an A earning 6 points, a B 5 points etc. The maximum of 2 points earned on the subsidiary and general paper are added to the points earned on the principal papers. This means the highest total score a subject could earn is 3 As and passes on the subsidiary and general paper, giving 20 points.

Standardised test scores will be constructed for each mock subject by subtracting the mean and dividing by the standard deviation of the control group. An overall index of mock exam performance will be calculated by summing test scores across all subjects and renormalising.

This means the equivalent mock score can be calculated for each of the S6 outcome variables:

1. **Total mock exam score**: Aggregate exam score composed of mock exam scores across all principal and subsidiary subjects taken by a student, with subsidiary subjects scored in points.
2. **Principal mock score**: Aggregate score composed of mock exam scores in the principal papers only.
3. **General paper and subsidiary paper mock score**: Standardised mock score on the general paper and subsidiary paper in maths or computer taken by all students. This will be an inverted scale of the 1-9 score on these papers.

3.1 Estimation and testing

I will estimate an ANCOVA specification of the following form:

$$y_{is1} = \beta_0 + \beta_1 \text{QofK} + y_{is0} + \boldsymbol{x_i'} \cdot \gamma + \theta_s + \epsilon_{is}, \tag{1}$$

where i indexes student at school s, y_{is1} denotes the exam outcome of interest, QofK is an indicator variable equal to one for if the student saw the movie The Queen of Katwe, $\boldsymbol{x_i'}$ is a vector of individual characteristics, θ_s is a vector of school fixed effects and ϵ_{is} is a random error.

y_{is0} is the mock exam result from before treatment. If available, the mock result in the specific outcome will be controlled for. If the equivalent mock result is not available for an outcome, the aggregate result constructed from the available mock papers will be controlled for instead.

This ANCOVA specification will be run for all outcomes detailed in this amendment and for all the outcomes in the original pre-analysis plan.

The parameter of interest is β_1, the treatment effects of the aspirational movie.

I include the following control variables in $\boldsymbol{x_i'}$ to improve precision, which were measured at the cinema during registration and confirmed with the schools.

1. whether the student is female
2. the age of the student
3. the number of subjects taken

For each outcome, I will test the following statistical hypotheses:

1. The aspirational movie has no effect on the outcome, $\beta_1 = 0$.

3.2 Subgroup Analysis

I will estimate heterogeneous treatment effects by augmenting equation 1 to include the variable and the interaction between treatment and the variable. The additional heterogeneity tests being added to the analysis are:

1. An indicator equal to one if the student was below the median exam performance in

their mock exams.

2. Indicators of decile of exam performance in the mock exam, with the excluded category being in the top decile.