Statistical Analysis Plan for "The effect of cash transfers on school re-enrollment during COVID-19 among vulnerable girls in informal settlements in Kenya: a randomized controlled trial"

Contents

1. Administrative information

Title of study	Cash transfers stimulate school re-enrollment among vulnerable adolescent girls in urban informal settlements as Kenya responds to COVID-19: a randomized controlled trial
Study registration number	ISRCTN 12792822 https://doi.org/10.1186/ISRCTN12792822
Study protocol version	Version 1 (20 November 2020)
SAP version	Version 1 (19 January 2021)
SAP revisions	N/A
Roles and responsibility	All designed the intervention and primary and secondary hypotheses. KA and BK are directing the intervention and data collection. ESH and JM will lead statistical analyses. All will interpret results and participate in manuscript preparation. KA – Karen Austrian (kaustrian@popcouncil.org) BK – Beth Kangwana (bkangwana@popcouncil.org) JM – John A. Maluccio (maluccio@middlebury.edu) ESH – Erica Soler-Hampejsek (erica.soler@gmail.com)

2. Introduction

2.1. Background and rational

After prolonged school closures starting in March and for most students lasting through December 2020[1] and extreme household economic stress due to the COVID-19 pandemic, there is substantial risk that adolescents will not return to school in 2021. In a related rapid survey covering part of the same sample we study, 88% of households reported having lost full or partial income due to the pandemic and 75% had skipped meals in prior week (Population Council 2020). Vulnerable adolescents who had been attending school in early 2020 are at high risk of not re-enrolling in school in 2021, leading to potential loss of key literacy and numeracy skills and lower final grade attainment. The survey indicated 87% of girls enrolled prior to school closings expected to return to school, but 57% worried that fees would be a barrier to returning. In Kenya, where there are significant bottlenecks transitioning to and completing secondary school, the implication for many could be non-completion of secondary school, increasing risk of adolescent pregnancy and losing a potentially useful credential in the labor market. The combination of household inability to pay school fees (which are typically due at the start of each of the three annual school terms), prioritization in households of education for boys and risk of pregnancy may have created an environment particularly hindering continued adolescent female education.

We present the research design and statistical analysis plan for a randomized controlled trial (Casey, Glennerster and Miguel 2012; Hiemstra et al. 2019). The study builds on a longitudinal cohort of girls first interviewed in 2015 (at ages 11–15) in informal settlements (Kibera and

[1] Closures in Kenya were among the longest in Africa (Reuters 2020). Some schools were partially reopened in October for grades 4, 8 and 12, three levels preparing for standardized exams rescheduled from November 2020 to March 2021.

Huruma) in Nairobi, Kenya and followed since as part of the Adolescent Girls Initiative-Kenya (AGI-K) study (Austrian et al. 2015, 2019, 2020; Kangwana et al. 2021).

2.2. Objectives and hypotheses

We evaluate the effect of a one-time cash transfer of ~US$150 on school re-enrollment for adolescent girls from Kibera and Huruma, two informal urban settlements in Nairobi, Kenya.

Primary Hypothesis: A one-time labeled cash transfer paid to the household when schools reopen after prolonged COVID-19 widespread closures will increase self-reported enrollment measured six weeks after the start of the first 2021 school term for targeted adolescent girls.

Secondary Hypotheses:

A one-time labeled cash transfer paid to the household when schools reopen after prolonged COVID-19 closures will:

II.1. Increase self-reported school enrollment at any point during the first six weeks after the start of the first 2021 school term (i.e., matriculation regardless of enrollment status measured at six weeks) for targeted adolescent girls.

II.2. Increase individual expectations or confidence that the targeted adolescent girls will complete secondary school.

II.3. Increase self-reported school enrollment measured six weeks after the start of the first 2021 school term for siblings (ages 6–20) of targeted adolescent girls.

II.4. Increase self-reported school enrollment measured six weeks after the start of the first 2021 school term for targeted adolescent girls who previously benefited (2015 – 2017) from the AGI-K education conditional cash transfer (CCT) program **more** than the increase for targeted girls who did not benefit from the prior CCT.

3. Study methods

3.1. Study design

Randomized controlled trial with one treatment and one control group, of equal size.

Cash transfer treatment arm: Households of targeted adolescent girls will receive a one-time cash transfer in early January 2021 (the week that schools reopen for all students). The transfer is KES 16,000 (~$150), paid into a bank account designated by the girl's parent or guardian.

Just after randomization in early December 2020 (3.2 Randomization), households were notified they would be receiving the transfer (after a baseline tracking survey, 5.2 Eligibility), with the accompanying SMS message (translated from Swahili):

> Thank you for participating in the recent survey. You have been randomly selected from our study participants to participate in the education programme of the study. Someone from Population Council will be calling you in the next few days to ask for your bank details. Kindly prepare to share this information with us - your bank account details including Name, Account Number and Branch Name. Thank You! Population Council.

When collecting bank details over the next week by phone, researchers shared the following message:

> [name of girl] has been randomly selected within the AGI-K cohort to benefit from a one-time education cash transfer. You will receive 16,000 KES in early January to support the cost of her schooling so that she can re-enroll when schools re-open after the COVID-19 closures. I would therefore wish to collect some bank account details that will facilitate in the disbursement of this cash transfer in January 2021.

Last, in early January (6–8) transfers were made and on January 11 the following message was sent.

> Happy New Year! We are delighted to inform you that you have received 16,000 KES in your account to support education for the AGI-K girl. Kindly check your bank account, and if there are any problems, feel free to contact [contact name at Population Council] on [contact phone number]. Population Council

Because of this accompanying messaging, we refer to it as a labelled cash transfer (Benhassine et al. 2015; Pace et al. 2019; Heinrich and Knowles et al. 2020).

Cash transfer control arm: No treatment (i.e., no cash transfer and no messaging at any point other than for the surveys)

An existing cohort of adolescent girls first interviewed in 2015 when they were ages 11–15 as part of AGI-K is targeted and the girls from all five AGI-K study arms shown in Table 1 were screened for program eligibility (Austrian et al. 2015, 2016, 2020). Specific eligibility criteria are detailed below (5.2 Eligibility). AGI-K randomized individual girls residing in Kibera to one of the first four study arms and also included a non-experimental comparison group from nearby Huruma.

Table 1. AGI-K interventions (2015–17)

AGI-K Intervention Package/Study Arm	Abbreviation
Violence Prevention Only (Kibera)	V-only
Violence Prevention + Education CCT (Kibera)	VE
Violence Prevention + Education CCT + Health (Kibera)	VEH
Violence Prevention + Education CCT + Health + Wealth Creation (Kibera)	VEHW
External Comparison site (Huruma)	Huruma

Participants were identified in early December 2020, approximately one month before schools were scheduled to reopen, and will be followed at six weeks after the start of the first 2021 calendar year school term (i.e., six weeks after the cash transfers are made). 50% of each AGI-K study arm was randomized into the cash transfer treatment and 50% into control (3.2 Randomization), with transfers made in early January 2021.

3.2. Randomization

From 24–29 November 2020, we carried out a tracking baseline survey by phone, attempting to contact N=1,912 eligible girls (5.2 Eligibility). Anticipating difficulties recontacting individuals by phone during the pandemic (e.g., due to individuals possibly no longer having phone service), we did not randomize the list of eligible girls prior to recontact but rather carried out randomization **after** verifying who could be contacted to participate in the study. This better ensured baseline balance after randomization, in particular ensuring equal numbers in treatment and control which is optimal for statistical power. N=1,620 (84.7%) girls, in 1,616 distinct households were successfully contacted and all invited to participate. Four households had two co-resident eligible girls each.

On 1 December 2020, we implemented Stata code to randomly allocate the cash transfer treatment and control groups. Invited girls were stratified by AGI-K study arm and randomized 50% to treatment and 50% to control at the individual girl level using Stata version 15 code:

```
set seed 12012020

sample 50, by(studyarm)

gen T=1

**Those not initially selected by "sample 50" are allocated to control**
```

In two of the four households with two co-resident girls, the above randomization assigned both girls to treatment. In the other two households, the above randomization assigned one girl to treatment and the other to control. We reassigned the latter two girls initially assigned as control to treatment, to ensure the same treatment status for both eligible girls within the household. These four households each receive a double (32,000 KES) transfer.

After this procedure, baseline random assignment was 813 girls to treatment and 807 to control.

Although there is some minimal risk of spillovers from treatment to control girls in the study in the densely populated study areas (Kibera had ~6500 households with over 36,000 people in 2015, more than 20,000 people per square kilometer; Huruma has similar population density), cluster or school-level randomization in this setting were not appropriate given ill-defined borders between "neighborhoods" in Kibera and Huruma (and the relatively small geographic area they cover) and school choice patterns (including both public and private options) in which girls rarely attend the nearest school often attending school several kilometers away making it possible for them to consider re-enrollment in potentially dozens of different schools (Maluccio et al. 2018).

3.3. Power considerations

The cash transfer is randomized to half of the girls in each of the original AGI-K study arms. The November 2020 baseline tracking survey indicated that 90% of the 1,620 girls interviewed were enrolled in school in March 2020. Intentions to return to school were measured in a related survey in June 2020 following a subset of the girls in the AGI-K study cohort and other similarly aged girls (Population Council 2020), and 87% of them expected to return to school. Thus, in the absence of any intervention, 78% of girls in the sample are expected to enroll in school in 2021. Given the sample size of 1,620 girls, a minimum detectable effects (MDE) approach was used to conduct power analysis using Stata 15.1 (StataCorp 2017). With power of 0.8, and significance level of 0.05, the study will be able to detect a minimum difference of 5.5 percentage points in school enrollment between girls randomized to receive the cash transfer and girls in the control arm. The study will also be able to detect a minimum difference of 0.14 standard deviations for secondary outcomes.

3.4. Hypothesis testing framework

For the primary outcome and secondary outcomes II.1–II.3, the null hypothesis is that there is no true difference in effect between the treatment and control arms. For secondary outcome II.4, the null hypothesis is that there is no true difference in the effect between treatment and control arms for girls in AGI-K study arms with and without the prior AGI-K CCT (subgroup analysis).

3.5. Statistical interim analysis and stopping guidance

N/A

3.6. Timing of final analysis

A follow-up survey is planned for six weeks after the start of the first 2021 school term (around mid-February 2021). A modest incentive paid in cell phone airtime (valued at KES 200) will be provided to respondents. Analysis is planned for March 2021.

3.7. Timing of outcome assessments

Primary and secondary outcomes measured six weeks after the start of the first 2021 school term (around mid-February 2021).

4. Statistical principles

4.1. Confidence intervals and *P*-values

All applicable statistical tests will be two-sided and at the 5% significance level, unless otherwise specified. Two-sided 95% confidence intervals will be presented whenever feasible. Robust standard errors will be estimated (StataCorp 2017).

4.2. Adherence and protocol deviations

Adherence is defined as verification of electronic delivery of the 16,000 KES cash transfer into the bank account of the parent or guardian of the target girl in the cash transfer treatment, for use by the household. Protocol deviations include non-delivery of the transfer into the bank account of the parent or guardian of a treatment girl or delivery of a transfer into the bank account of the parent or guardian of a control girl.

4.3. Analysis populations

Analysis will be carried out on all respondents in the follow-up survey February 2021 with valid outcome information, separately for each outcome and on an intent-to-treat basis.

5. Study population

5.1. Screening data

See 5.2 Eligibility

5.2. Eligibility

The initial sample frame is based on the AGI-K cohort first interviewed in 2015 (Austrian et al. 2015) and followed up in 2017 and 2019 (Austrian et al. 2019, 2020). We note that approximately 20% of this sample has also been followed via rapid phone surveys related to the pandemic throughout 2020 (Population Council 2020). The sample was screened based on the following criteria.

Inclusion criteria:

1) Interviewed in 2019 as part of AGI-K cohort
2) Interviewed in 2019 and had not completed secondary school
3) Interviewed in 2019 and was not enrolled in final year of secondary school
4) Interviewed in 2019 and had enrolled in school at some point between 2017–19

Exclusion criteria:

1) Not interviewed in 2019 AGI-K follow-up survey
2) Interviewed in 2019 and had completed secondary school
3) Interviewed in 2019 and was enrolled in final year of secondary school
4) Interviewed in 2019 and had not enrolled in school between 2017–19

These criteria yielded 1,912 eligible girls. Table 2 indicates by AGI-K study arm the exclusions for each reason.

Table 2. Determination of eligibility

Determination of eligibility							
				AGI-K Study arm			
		V-only	VE	VEH	VEHW	Huruma	Total
2015 AGI-K baseline	N	597	592	609	592	662	3,052
Interviewed in 2019	N	482	528	533	532	483	2,558
	% baseline	80.74	89.19	87.52	89.86	72.96	83.81
Completed Form 4	N	20	21	22	18	19	100
In Form 4 in 2019	N	82	117	105	102	37	443
Not in school 2017-19	N	22	14	13	12	42	103
Eligible	**N**	**358**	**376**	**393**	**400**	**385**	**1,912**
	% baseline	59.97	63.51	64.53	67.57	58.16	62.65

From 24–29 November 2020, we carried out a tracking baseline phone survey attempting contact of all N=1,912 eligible girls. N=1,620 (84.7%) were contacted and all invited to participate in the study. In addition to updating parent or guardian contact and location information, the brief survey included questions about the girl's current school plans, fertility and civil status, and background household characteristics to be used in the construction of a wealth index score described below and estimated for in previous work (Filmer and Pritchett 2001; Maluccio et al. 2018; Austrian et al. 2020).

Table 3 indicates by AGI-K study arm the contact rates and Table 4 provides a summary of reasons for non-interview. Non-interview was primarily due to being unable to contact the girl (nearly always because the call did not go through) but also due to refusals, higher in the V-only and Huruma external comparison AGI-K study arms. Other includes having moved to an unknown location; one girl moved to Tanzania; and one girl was deceased.

Table 3. Tracking baseline survey interview and invitation to participate in study

Tracking baseline survey interview and invitation to participate in study							
				AGI-K Study arm			
		V-only	VE	VEH	VEHW	Huruma	Total
Total eligible	N	358	376	393	400	385	1,912
Not interviewed	N	73	38	39	47	95	292
	%	20.4	10.1	9.9	11.8	24.7	15.3
Interviewed and invited to participate	N	285	338	354	353	290	1,620
	%	79.6	89.9	90.1	88.3	75.3	84.7

Table 4. Tracking baseline survey reasons for non-interview

Tracking baseline survey reasons for non-interview							
				AGI-K Study arm			
		V-only	VE	VEH	VEHW	Huruma	Total
Not interviewed	N	73	38	39	47	95	292
Unable to contact	N	58	35	31	40	68	232
	%	79.5	92.1	79.5	85.1	71.6	79.5
Refusal	N	14	3	7	7	26	57
	%	19.2	7.9	17.9	14.9	27.4	19.5
Other	N	1	0	1	0	1	3
	%	1.4	0.0	2.6	0.0	1.1	1.0

After initial contact, invited girls were randomly assigned to treatment or control, stratified by AGI-K study arm (3.2 Randomization). The resulting samples from randomization are shown in Table 5, with 813 girls randomized to cash transfer treatment and 807 to control. At this point in early December 2020, girls randomized to treatment were notified about the upcoming transfer and bank account information collected as detailed above (3.1 Study Design).

Table 5. Random allocation to cash transfer treatment and control

Random allocation to cash transfer treatment and control							
				AGI-K Study arm			
		V-only	VE	VEH	VEHW	Huruma	Total
Control	N	140	169	177	176	145	807
	%	49.1	50.0	50.0	49.9	50.0	49.8
Treatment	N	145	169	177	177	145	813
	%	50.9	50.0	50.0	50.1	50.0	50.2
Total	N	285	338	354	353	290	1,620
	%	100.00	100.00	100.00	100.00	100.00	100.00

5.3. Recruitment

As detailed in 5.2 Eligibility, in the baseline tracking survey we attempted contact of all N=1,912 eligible adolescent girls and successfully interviewed N=1,620 (84.7%), all of whom were invited to participate in the program and randomized to cash transfer treatment or control.

In addition to details provided in this analysis plan, the sample flow or CONSORT diagram will report receipt of the cash transfer (whether deposited to the bank account of the parent or guardian in early January) and mid-February follow-up survey result including reasons for non-interview as necessary. See Table 6 for the template; results by AGI-K study arm will be reported in an appendix.

Table 6. CONSORT diagram template

CONSORT TEMPLATE					
Attempted contact of 1,912 eligible girls					
Successful contact of 1,620 (84.7%) in baseline tracking survey					
- All invited to participate					
Randomization					
	Treatment		Control		
	N=813		N=807		
Transfer delivered					
January 6-8	N=800		N=0		
Januar 15-xx	N=802		N=0 (expected zero)		
Total	N=###		N=0 (expected zero)		
February follow-up survey					
Interviewed	N=###		N=###		
(%)	##.#		##.#		
Loss to follow-up	N=###		N=###		
Reason 1					
Reason 2					
Etc					

Transfers in the amount of KES 16,000 were delivered January 6th to 8th to the bank accounts of the parents/guardians of 800 girls assigned to treatment. Initially, 13 households assigned to treatment could not be contacted or refused receipt of the transfer; they were given until January 15 and at least 2 of the 13 came forward to accept the transfer by that date. The final CONSORT diagram will report all transfers delivered out of the 813, expected to be 802 or more. No transfers were made to control girls.

5.4. Withdrawal/follow-up

We will report delivery of the intervention (cash transfers made, see 4.2 Adherence and 5.3 Recruitment) and loss to follow-up by cash transfer treatment and control, including reasons for loss to follow-up where available (see 5.3 Recruitment).

5.5. Baseline characteristics

The following table presents baseline characteristics by randomized treatment arm for the sample invited to participate in the program (N=1,620) using measurements from: 1) the AGI-K cohort 2015 baseline survey; 2) the AGI-K four-year follow-up 2019 survey; and 3) the tracking baseline survey completed in November 2020. We present standardized differences (the difference between treatment and control divided by the overall standard deviation of the variable) but do not calculate statistical tests for these comparisons as recommended in CONSORT guidelines for randomized trials (Moher et al. 2010). Unsurprisingly given randomization, means across treatment and control are very similar. All but one of the

reported standardized differences are less than 0.1 standard deviation. Such small differences were expected given the randomized design.

Table 7 panel A presents baseline 2015 characteristics for the N=1,620 sample interviewed in November 2020 (for eventual comparison using the same variables on the sample of observations successfully interviewed in February 2021).

Table 7 panel B presents time-varying variables collected in 2015, 2019 and November 2020 using all available observations for the sample of N=1,620. These include quintiles for the wealth index described in Austrian et al. (2020) calculated in 2015 for the 2015 and 2019 columns and re-estimated using November 2020 baseline tracking survey data for the 2020 column. Below the wealth index we present its constituent elements. For 2015, the wealth index, which is the first component from the principal components analysis of the elements explained 25.6% of the variation. In 2020 the percent of variation explained was similar—26.5%.

We will present the same sets of comparisons for the (non-attrited) follow-up sample of girls interviewed in February 2021 as indicated in the tables. Because attrition between November 2020 and February 2021 is possibly non-random, however, for those comparisons we will also test for significant differences between treatment and control using standard t-tests controlling for stratification for each outcome and using robust standard errors.

6. Analysis

6.1. Outcome definitions

Primary outcome: Self-reported enrollment measured six weeks after the start of the first 2021 calendar year school term for targeted adolescent girls. Constructed from responses to "Is [girl name] currently enrolled in school?"

Secondary outcomes:

II.1. Self-reported school enrollment at any point during the first six weeks after the start of the first 2021 school term (i.e., matriculation regardless of enrollment status measured at six weeks) for targeted adolescent girls. Constructed from responses to "Did [girl name] enroll in school this year, even if she is no longer attending?" asked if girl is not currently enrolled.

II.3. Individual-level expectations targeted adolescent girls will complete secondary school, constructed from responses to the following question (into a binary variable indicating high likelihood). "What are the chances that [girl name] will finish secondary school? Would you say high, about 50-50 or low?".

II.3. Self-reported school enrollment measured six weeks after the start of the first 2021 school term for siblings ages 6–20 of targeted adolescent girls. Measured as the fraction of enrolled siblings to total siblings (6–20) and, separately, by gender. Constructed from responses to the following three questions (for each gender):

"How many sisters do you have between the ages of 6 and 20?"

"How many of them were enrolled in school in March 2020, just before COVID-19 started?"

"How many of them are currently enrolled in school?"

II.4. Self-reported school enrollment measured six weeks after the start of the first 2021 school term for targeted adolescent girls who previously benefited (2015–17) from the AGI-K

education conditional cash transfer (CCT) program **more** than the increase for targeted girls who did not benefit from the prior CCT. See primary outcome measure above.

Note: The adolescent girl is the target for interview and all questions will be framed as second person "you" unless it is only possible to interview the guardian in which case in third person detailed above. We will report the fraction of self- versus third-person reports.

6.2. Analysis methods

MAIN ANALYSES

The main ITT effect of the one-time cash transfer on the primary outcome and secondary outcomes II.1 and II.2 will be estimated using:

$$(1) \dots Y_{is} = \beta_0 + \beta_1 T_i + \boldsymbol{\alpha_s} + \boldsymbol{age_i} + \beta_2 educ_i + \varepsilon_{is}$$

where

Y_{is} is the outcome for individual i in AGI-K study arm s;

T_i equals 1 if individual (or household) is assigned to cash transfer treatment and 0 otherwise;

$\boldsymbol{\alpha_s}$ are stratification fixed effects (one per AGI-K study arm omitting V-only) per the randomized design;

$\boldsymbol{age_i}$ are binary indicators for age in years at AGI-K start in 2015, omitting the oldest, age 15);

$educ_i$ is the 2019 highest grade completed for individual *i.*

ε_{is} is an assumed idiosyncratic error term.

In addition to the binary indicators for the AGI-K study arms (omitting V-only) used in the stratified randomization, given the relationships between current schooling outcomes and expectations with age and prior schooling level, we control for them in the main specification. (We note that including highest grade completed is similar to estimating ANCOVA, but since the study sample was selected from girls enrolled in school in 2019 prior enrollment does not vary in the sample.)

β_1 yields the estimated ITT effect of the cash transfer treatment on outcome Y_{is}.

To evaluate II.3, we modify equation (1) by excluding the target girl-specific age and education controls, including instead the numbers of female and male siblings ages 6–20.

SENSITIVITY ANALYSES

To assess internal validity of the experimental results, we will conduct three sensitivity analyses related to attrition for the primary and secondary outcomes II.1–II.3.

As described above, there were N=1,912 girls eligible for the cash transfer program and in November 2020 we successfully contacted and invited to participate N=1,620 (5.2 Eligibility). Attrition in part reflects the challenges of longitudinal survey work during the COVID-19 pandemic when there has been substantial economic and other disruption, including residential mobility. Although the relatively short 3-month time frame between the end-November baseline and mid-February follow-up will likely make tracing easier, we nevertheless expect there will be some attrition in the February 2021 follow-up survey. This will reduce sample sizes and depending on the nature of attrition across cash transfer

treatment and control groups potentially threaten the internal validity of the randomized experiment (Molina Millán and Macours 2017).

The first sensitivity analysis incorporates additional controls to increase precision and reweights for attrition. We will follow Austrian et al. (2020) and incorporate a limited set of additional covariates measured in 2015, shown in Table 7 panel A. They include the girl's cognitive score in 2015, whether her parents had completed primary school and whether she lived with both parents.

We will also incorporate any variable from the November 2020 tracking baseline survey showing an imbalance after attrition of more than a 0.10 standardized difference as determined via the comparison of baseline characteristics for the 2021 sample (5.5 Baseline Characteristics). 0.10 is chosen based on results in Table 7 panels A and B, where only one comparison was larger (household owns television in 2019) after randomization.

In addition to including these controls, we will use inverse probability weighting (IPW) estimating the probability of re-interview in February 2021 for the N=1,620 sample, using all of the variables shown in Table 7 and following the methodology implemented in Austrian et al. (2020) outlined below. In addition, we will use an indicator variable for whether the girl has been included in the on-going 2020 rapid phone survey sample, since renewed contact with the girl prior to November 2020 may increase the probability of re-interview. In June 2020, 333 (20.6%) of the sample were interviewed as part of that on-going survey.

We will calculate inverse probability weights (IPW) for girl as follows. First, we will impute the small number of missing values on any variable in Table 7 panels A and B using the median value from within the relevant AGI-K study arm (for example, the 10 missing baseline cognitive scores in panel A). Second, we will estimate bivariate regressions on an indicator of being interviewed at the February 2021 follow-up for each variable (separately for cash transfer treatment and control groups) on the sample of participants (N=1,620). We retain for potential inclusion in the weight construction all variables significant at 10%. Third, each retained variable, along with indicators for age in years and the AGI-K strata indicators, is directly included and also interacted with a binary indicator of cash transfer treatment so that the weights are calculated separately for each treatment group. Fourth, we estimate the probability of being interviewed on this set of baseline predictors. To account for collinearity between predictors, the baseline predictor set is further limited by conducting stepwise selection of variables with backward elimination and using the adjusted R^2 as the information criteria. Indicators for treatment and age are fixed in the regressions. At each step, the iterative procedure removes from the model the predictor that most improves the information criterion until there is no variable whose removal improves it. We implement this using 'vselect' in Stata (Lindsey and Sheather 2010). Fifth, using the final model, we predict for each observation the probability of having been re-interviewed and construct the IPW. We will also report information about the distribution of the constructed weights.

The second sensitivity analysis will estimate Lee bounds for attrition based on equation (1) specifications and tightening using the variables described in equation (1). Results will not be reported if the monotonicity condition is violated (Lee 2009; StataCorp 2017).

The third sensitivity analysis will exclude the Huruma external comparison, the only non-experimental AGI-K study arm for which attrition was highest throughout the cohort study.

SUBGROUP ANALYSES (including Hypothesis II.4)

We will estimate two different subgroup analyses to consider heterogeneous effects.

The first subgroup analysis is for the primary outcome and secondary outcomes II.1 and II.2

We will explore in a single equation whether program impacts were heterogeneous (Baird et al. 2020) with respect to two important observable characteristics associated with continued schooling: completed grades (most recently collected in 2019 so for end-year 2018) of the girl and economic status of the household. To estimate we will use:

$$(2) \dots Y_{is} = \beta_0 + \beta_1 T_i + \boldsymbol{\alpha_s} + \boldsymbol{age_i} + \beta_2 Form2_i + \beta_3 T_i \times Form2_i + \beta_4 HiSES_i + \beta_5 T_i \times HiSES_i + \varepsilon_{is}$$

Where

$Form2_i$ equals 1 if the girl had completed form 2 (grade 10) or higher by 2019 and 0 otherwise;

We constructed a socioeconomic status index constructed from the first principal component from a principal components analysis of a set of assets and housing characteristics used in previous work (Maluccio et al. 2018; Austrian et al. 2020) and based on variables from the November 2020 survey. This variable is shown in the baseline comparison Table 7 panel B, categorized by quintile. For the regression analysis it will be dichotomized into low (lower two quintiles) and high SES.

For these specifications we will combine β_1, β_3, and β_5 to determine the estimated effects for the four subgroups and test differences across them.

	< Form 2 completed in 2019	>= Form 2 completed in 2019
Low SES	β_1	$\beta_1 + \beta_3$
High SES	$\beta_1 + \beta_5$	$\beta_1 + \beta_3 + \beta_5$

Subgroup analyses are to understand whether some important subgroups were affected differently, to improve potential targeting and to shed light on possible underlying mechanisms of any observed effects.

We will also explore possible mechanisms by examining the subsample of those not re-enrolled in school in February 2021 and comparing the reported reasons for not having enrolled across treatment groups (as fractions of the total number not in school in treatment and control).

The second subgroup analysis is to evaluate hypothesis II.4 for which we will use:

$$(3) \dots Y_{is} = \beta_0 + \beta_1 T_i + \beta_2 T_i \times CCT_i + \boldsymbol{\alpha_s} + \boldsymbol{age_i} + \beta_3 educ_i + \varepsilon_{is}$$

where all variables are the same as above except

CCT_i equals 1 if the individual was assigned the AGI-K education conditional cash transfer (from 2015–17, AGI-K study arm VE, VEH or VEHW) and 0 otherwise.

$educ_i$ is the **2015** highest grade completed for individual *i* (prior to the start of the AGI-K CCT).

β_1 yields the estimated ITT effect of the cash transfer treatment on Y_{is} for girls not previously exposed to the AGI-K CCT and β_2 yields the differential effect of the cash transfer treatment on Y_{is} for individuals who were previously exposed (so that $\beta_1 + \beta 2_2$ represent the total effect for those girls with prior exposure to the AGI-K CCT).

While not all individuals from the original AGI-K cohorts were followed for the cash transfer study, we note that by virtue of the original AGI-K randomization the stratification indicators for VE, VEH and VEHW (all relative to V-only) in $\boldsymbol{\alpha_s}$ provide estimates of the effect of the AGI-K study arms net of the one-time 2021 cash transfer, after nearly six years. In this analysis we condition on highest grade in the 2015 baseline rather than in 2019 (as in equation 1). Although there were only modest effects on grade completion resulting from the AGI-K intervention and no significant effects on secondary school completion (Kangwana et al. 2021), the baseline control is used so that long-term effects can be estimated without them operating through "current" or 2019 grades completed if it were in the equation. To ensure identification is more directly linked to the randomized designs, the models will be estimated excluding Huruma, the non-experimental AGI-K study arm.

6.3. Missing data

Analyses will be carried out on samples with valid outcome data, using all available observations for each outcome. Missing data for variables used in regression adjustment or IPW weight construction will be imputed using the median value for the specific AGI-K study arm and cash transfer study treatment group.

6.4. Additional analyses

Using data from November 2020, we carried out principal components analysis on the variables indicated as elements of the wealth index in Table 7, Panel B. This common approach for generating a wealth index (Filmer and Pritchett 2001) has been used throughout the AGI-K cohort study. We retain the first principal component as an index measure of wealth. Results of the analysis are shown below and indicate the first principal component explained 26.5% of the variation in the set of variables. All elements had expected loadings (note that hungry is negative) on the first principal component 0.25 or higher and 5 of 10 loaded 0.3 or higher.

```
Principal components/correlation                Number of obs    =      1,620
                                                Number of comp.  =          3
                                                Trace            =         10
    Rotation: (unrotated = principal)           Rho              =     0.5078
```

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	2.64509	1.22107	0.2645	0.2645
Comp2	1.42402	.414916	0.1424	0.4069
Comp3	1.0091	.0894051	0.1009	0.5078
Comp4	.919698	.0236731	0.0920	0.5998
Comp5	.896025	.118387	0.0896	0.6894
Comp6	.777637	.0424719	0.0778	0.7672
Comp7	.735165	.0731023	0.0735	0.8407
Comp8	.662063	.160283	0.0662	0.9069
Comp9	.50178	.0723586	0.0502	0.9571
Comp10	.429421	.	0.0429	1.0000

Principal components (eigenvectors)

Variable	Comp1	Comp2	Comp3	Unexplained
livestock	0.2836	0.5736	0.1200	.3042
tv	0.2446	-0.3836	-0.1766	.6007
watch	0.2877	-0.2088	-0.1834	.6851
mosnet	0.2611	0.1680	-0.4467	.5781
sleepingroom	0.3037	0.0330	-0.4824	.5196
agrland	0.2320	0.5959	0.1470	.3301
hungry	-0.3082	-0.0092	0.3494	.6254
sav1000	0.3739	-0.2105	0.2991	.4768
sav5000	0.4430	-0.2075	0.3450	.2995
sav10000	0.3628	-0.0914	0.3693	.5023

6.5. Harms

N/A

6.6. Statistical software

Statistical analysis will be undertaken using Stata 15.0 version or later.

References

Austrian K, Muthengi E, Mumah J, Soler-Hampejsek E, Kabiru CW, Abuya B, Maluccio JA: The adolescent girls initiative-Kenya (AGI-K): study protocol. *BMC public health* 2016, 16(1):210.

Austrian K, Muthengi E, Riley T, Mumah J, Kabiru C, Abuya B: Adolescent Girls Initiative-Kenya. Baseline report. 2015.

Austrian, K., E. Soler-Hampejsek, B. Kangwana, Y. Dibaba Wado, B. Abuya and J.A. Maluccio. 2020. Impacts of two-year multi-sectoral interventions on young adolescent girls' education, health and economic outcomes: Adolescent Girls Initiative-Kenya randomized trial. Population Council, Nairobi, Kenya.

Baird, S., B. Ozler, Ch. Dell`Aira and D.U. Salam. 2020. Using group interpersonal psychotherapy to improve well-being of adolescent girls. *Journal of Development Economics Registered Report Stage 1*.

Benhassine, N., F. Devoto, E. Duflo, P. Dupas, and V. Pouliquen. 2015. Turning a shove into a nudge? A "labeled cash transfer" for education. American Economic Journal: Economic Policy, 7(3): 86–125.

Casey, K., R. Glennerster and E. Miguel. 2012. Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *Quarterly Journal of Economics*, 1755–1812.

Filmer, D. and L. Pritchett. 2001. Estimating wealth effects without expenditure data – or tears: An application to educational enrollments in states of India. *Demography*, 38: 115– 32.

Heinrich, C. and M.T. Knowles. 2020. A fine predicament: Conditioning, compliance and condequences in a labeled cash transfer program. World Development, 129 (May) 104876.

Hiemstra, B., F. Keus, J. Wetterslev, C. Gluud, I.C.C. van der Horst. 2019. DEBATE-statistical analysis plans for observational studies. *BMC Medical Research Methodology*, (2019) 19: 233.

Kangwana, E., K. Austrian, J.A. Maluccio, E. Soler-Hampejsek et al. 2021. Impact of a Cash Transfer and Girls Empowerment Program on Early Sexual Debut and Fertility in a Kenyan Urban Informal Settlement: Results from a Mixed Methods Randomized Trial. Photocopy, Population Council-Nairobi Kenya.

Lee, D.S. 2009. "Training, Wagers, and Sample Selection: Estimating Sharp Bounds on Treatment Effect." *The Review of Economic* Studies 76(3):1071–1102.

Lindsey, C. and S. Sheather. (2010). "Variable selection in linear regression." *The Stata Journal*, 10(4): 650–669.

Maluccio, J.A., M. Hussein, B.A. Abuya, E. Muluve, E. Muthengi and K. Austrian. 2018. Adolescent girls primary school mobility and educational outcomes in urban Kenya. *International Journal of Educational Development*, 62(September): 75–87.

Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol 2010;63: e1–37.

Molina Millán, T. and K. Macours. 2017. "Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias." *CEPR Discussion Paper No. DP11962*

Pace, N., S. Daidone, B. Davis and L. Pellerano. 2019. Shaping cash transfer impacts through 'soft-conditions': Evidence from Lesotho. Journal of African Economics, 28(1): 39–69.

Population Council. 2020. Kenya: COVID-19 Perceptions, Prevention Practices, and Impact: Responses from fourth round of data collection in five Nairobi informal settlements (Kibera, Huruma, Kariobangi, Dandora, and Mathare). Nairobi, Kenya.

Reuters. https://www.reuters.com/article/healthcoronavirus-kenya-schools/kenya-partially-reopens-schools-six-months-after-covid-closed-them-idUSL8N2H106G

StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.

Table 7. Comparison of means by treatment and control group								
Panel A: Baseline 2015 characteristics								
		2020				2021		
	T	C	STD		T	C	STD	p-value
Baseline age in 2015, years	12.19	12.20	-0.007		*To be completed using using same 2015*			
(sd)	1.13	1.18			*data, limited to final Feb 2021 sample*			
	813	807						
Cognitive score (0–16), mean	8.16	7.94	0.072					
(sd)	3.12	3.08						
	808	802						
Mother completed primary school =1	0.60	0.62	-0.041					
(sd)	0.49	0.48						
Father completed primary school = 1	0.73	0.75	-0.033					
(sd)	0.44	0.43						
Lived with both parents in 2015 = 1	0.54	0.55	-0.021					
(sd)	0.50	0.50						
	810	802						
All data measured in AGI-K 2015 baseline survey shown for sample interviewed in November 2020 baseline tracking survey								
P-value test of significance of differences in means across treatments controlling for stratification (5.5 Baseline)								

Table 7. Comparison of means by treatment and control group														
Panel B: Time-varying characteristics														
		2015			2019			2020			2021			
	T	C	STD	T	C	STD	T	C	STD	T	C	STD	p-value	
Enrolled in current school year = 1	1.00	1.00	-0.028	0.92	0.91	0.051	0.92	0.89	0.094	*To be completed using using updated*				
	0.05	0.04		0.27	0.29		0.28	0.31		*responses from final Feb 2021 sample*				
	813	807		811	803		813	807						
Current grade (if enrolled)	6.16	6.17	-0.005	9.90	9.85	0.046	10.87	10.85	0.024					
	1.12	1.14		1.07	1.07		1.17	1.25						
	811	805		714	685		746	718						
Grade attainment, mean	5.18	5.20	-0.020	8.92	8.86	0.049	n.a.							
(sd)	1.13	1.13		1.11	1.11									
	813	807		813	807									
Primary school completion = 1	0.01	0.02	-0.051	0.90	0.89	0.026	n.a.							
	0.11	0.14		0.30	0.31									
	813	807		813	807									
Currently married = 1	0.000	0.002	-0.071	0.02	0.03	-0.058	0.01	0.02	-0.068					
	0.000	0.050		0.14	0.17		0.09	0.13						
	813	807					813	807						
Have child = 1	0.000	0.001	-0.050	0.03	0.04	-0.081	0.05	0.07	-0.099					
	0.000	0.035		0.17	0.20		0.22	0.26						
	813	807		813	806		813	807						
Currently pregnant = 1	0.00	0.00	0.000	0.01	0.01	0.046	0.01	0.02	-0.065					
	0.00	0.00		0.12	0.09		0.12	0.15						
	813	806		813	807		810	802						
Household wealth quintile (1-5), mean	3.00	3.06	-0.037	3.17	3.23	-0.044	2.99	3.00	-0.010	*To be completed using using same 2015*				
	1.43	1.45		1.48	1.47		1.41	1.42		*data, limited to final Feb 2021 sample*				
	810	802		812	807		813	807						
Elements of wealth index														
HH owns any livestock = 1	0.28	0.27	0.025	0.22	0.22	0.018	0.20	0.20	-0.007					
	0.45	0.44		0.42	0.41		0.40	0.40						
	810	802		812	807		813	807						
HH owns television = 1	0.74	0.79	-0.110	0.73	0.72	0.027	0.70	0.70	-0.001					
	0.44	0.41		0.44	0.45		0.46	0.46						
	813	807		813	807		813	807						
HH owns clock or watch = 1	0.32	0.30	0.027	0.28	0.27	0.034	0.25	0.25	0.013					
	0.47	0.46		0.45	0.44		0.44	0.43						
	813	807		813	807		813	807						
HH owns mosquito net = 1	0.39	0.38	0.022	0.40	0.42	-0.041	0.37	0.40	-0.067					
	0.49	0.48		0.49	0.49		0.48	0.49						
	813	807		813	807		813	807						
Number of rooms in HH for sleeping, mean	1.34	1.35	-0.024	1.51	1.54	-0.035	1.41	1.42	-0.022					
	0.62	0.61		0.70	0.70		0.67	0.63						
	810	802		812	807		813	807						
HH owns agricultural land = 1	0.41	0.41	-0.011	0.35	0.38	-0.067	0.31	0.32	-0.026					
	0.49	0.49		0.48	0.49		0.46	0.47						
	813	807		813	807		813	807						
HH went without food 1 day last month = 1	0.52	0.52	-0.015	0.48	0.46	0.044	0.74	0.72	0.041					
	0.50	0.50		0.50	0.50		0.44	0.45						
	810	802		812	807		813	807						
HH has enough savings or something to sell if need:														
KES 1,000 = 1	0.54	0.54	-0.018	0.61	0.63	-0.033	0.52	0.49	0.059					
	0.50	0.50		0.49	0.48		0.50	0.50						
	810	802		812	807		813	807						
KES 5,000 = 1	0.24	0.27	-0.080	0.31	0.34	-0.057	0.17	0.19	-0.058					
	0.42	0.44		0.46	0.47		0.38	0.40						
	810	802		812	807		813	807						
KES 10,000 = 1	0.07	0.10	-0.079	0.15	0.15	0.015	0.04	0.06	-0.065					
	0.26	0.29		0.36	0.35		0.20	0.23						
	810	802		812	807		813	807						
All available observations for sample of individuals invited to participate (N=1,620)														
Household wealth index is the first principal compoment from principal components analysis of the indicated items calcuated in 2015 and 2020														
P-value test of significance of differences in means across treatments controlling for stratification (5.5 Baseline)														