Wheels of Change: Transforming Girl's Lives with Bicycles *

Nathan Fiala† Ana Garcia-Hernandez‡ Kritika Narula§ Nishith Prakash¶

March 8, 2021

Abstract

We study the impact of a program that provides a bicycle to a school-going girl who lives more than 3 km from the school. We randomized whether a girl receives a bicycle with a small cost to her family to cover replacement parts, a bicycle where these costs are covered by the program and so is zero cost to the family or a control group. We find that the bicycle reduced average commuting time to school by 35%, late arrival by 66%, and decreased absenteeism by 27% in the short and medium-run. We also find evidence of increased grade transition in the medium-run, improved math test scores, girls expressing higher feelings of control over their lives and, for those who received bicycles with a small cost to her family, higher levels of aspirations, self-image, and a desire to delay marriage and pregnancy. Heterogeneity analysis by distance to school shows an inverted U-shape for most of the schooling and empowerment results, suggesting those impacts are larger for girls that live further away from school. This also suggests that empowerment outcomes worked through increased attendance in school.

Keywords: Girls' Education, Attendance, Test Scores, Bicycles, Female Aspiration, Female Empowerment, Safety, Zambia.

JEL Codes: H42, I21, I25, J16, O15.

*We are grateful for comments provided at various points during this study by Prashant Bharadwaj, Lasse Brune, Alex Cohen, Alexander Coutts, Manisha Desai, Taryn Dinkelman, Esther Duflo, Thomas Fujiwara, Siddharth George, Rachel Glennerster, Rema Hanna, Macartan Humphreys, Lakshmi Iyer, Tarun Jain, Dean Karlan, Daniel Keniston, Asim Khwaja, Michael Kremer, Matt Lowe, Nirajana Mishra, Mushfiq Mobarak, Dilip Mookherjee, Priya Mukherjee, Karthik Muralidharan, Nathan Nunn, Daniele Paserman, Gautam Rao, Deepak Saraswat, Alexandra Scacco, Frank Schilbach, Christopher Udry, Anjali Verma and seminar and conference participants at 3^{rd} Annual Economics Conference (Ashoka University), 15^{th} Annual Conference on Economic Growth and Development, Indian Institute of Management Ahmedabad, University of Rochester, Boston University, Emory University, Harvard Kennedy School (WAPPP), University of Notre Dame, Reserve Bank of India, Louisiana State University, University of Wisconsin-Madison, University of Wisconsin-La Crosse, International Food Policy Research Institute, and The World Bank. We thank Mahesh Dahal, Margarida Ortigão, and Olivia Siegl for excellent research assistance and the study participants for generously giving their time. We thank Emma Lambert-Porter, the field officers at Innovations for Poverty Action Zambia and all the staff from World Bicycle Relief Zambia, especially Alisha Meyers, and Lawrence Banda. The study was pre-registered with the AEA Trial Registry (RCT ID: AEARCTR-0003339). Funding for this study was provided by Innovations for Poverty Action, World Bicycle Relief, UBS Optimus Foundation, and University of Connecticut. Any errors are entirely our own.

†University of Connecticut, Makerere University and RWI-Leibniz Institute for Economic Research. nathan.fiala@uconn.edu.

‡Universidad del Rosario and Innovations for Poverty Action. ana.garciah@urosario.edu.co.

§Yale University. kritika.narula@yale.edu.

¶University of Connecticut, IZA, GLO, HiCN, and CReAM. nishith.prakash@uconn.edu.

1 Introduction

In the last two decades, developing countries have made considerable progress in closing the gender gap in education: women have more education today than they did fifty years ago in every country in the world (Barro and Lee (2013)). Despite this, adult women have less education than men in more than two-thirds of the world's countries (Evans et al. (2020)). There are several binding constraints for girls around the world, including the cost of schooling, social norms of schooling, safety, and distance to school. This is especially true in Zambia, where girls need to walk long distances to school and face harassment along the way (DHS (2002)). Starting in higher grades of primary school, adolescent girls are more likely to drop out of school than their male counterparts (Bank (2018)). Though long-distances to school is a problem for everyone, it is especially critical for girls who reach puberty around this age and face additional risks, including the danger of being assaulted on the way to school (Hawke (2015)). *In this paper, we evaluate the impact of a conditional in-kind transfer program that provides a bicycle to a school-going girl on measures of education and empowerment in rural Zambia.* Reducing gender disparities in education is an important policy question for developing countries. In addition to being one of the United Nations Sustainable Development Goals in 2030, it is also considered to be an important pathway to women's empowerment (Lundberg and Pollak (1993)). Although Zambia has reduced the gender gap in primary schooling dramatically over the past decade, gaps still exist in school enrollment, attendance, and dropouts, and this gap increases with age.[1] Moreover, parity in education access has not translated to parity in effective education attainment and learning (Saito (2011)). One of the main reasons cited by researchers is the long distance to school, which affects girls differentially than boys (Hawke (2015)), and is a larger problem in rural areas compared to urban areas (DHS (2002)).[2]

Almost 98% of the students in our sample walk to school, and on an average they travel approximately 110 minutes one way. Walking such long-distances to school can affect the intensive margin of learning through its impact on student *absenteeism*, *punctuality* and *tiredness*. National Assessment results in Zambia show that the longer the distance pupils traveled to school, the lower their learning achievement UNICEF (2014). In so far as gender parity in human capital acquisition is necessary for long-term growth and a high standard of living (Duflo (2012)), a policy aimed at improving girls'

[1] For example, starting in grade 6, significantly more girls leave school than boys. In 2013-14 girls dropped out of school more than double the rate of boys in grade 7 and three times the rate of boys in grade 11 (McConnell and Mupuwaliywa (2015)).

[2] The low population density in rural Zambia presents a challenges of access to school because homes are isolated from each other, causing children to walk more than 5 km to school UNICEF (2014).

effective education attainment is important for a country like Zambia, with one of the highest levels of gender inequality in the world.[3]

We study the impact of providing bicycles to adolescent girls in rural Zambia by experimentally varying the distribution of bicycles to girls in school. We hypothesize that the provision of bicycles will have an impact on their *educational attainment* and *empowerment*. First, by decreasing the daily distance cost of school attendance, we expect a *direct* impact on access to education.[4] Further, we expect bicycles to provide a safer way of travelling to school, which could also improve girls' access to education.[5] Second, the provision of bicycles could *directly* improve empowerment outcomes for girls. Scholars of the women's suffrage movement, including Susan B. Anthony in the United States, have highlighted the role that bicycles played in empowering women in the 19^{th} century by enhancing their independence, control, self-reliance, self-respect, mobility, and freedom (Macy (2011)). Improvement in educational attainment could also feed into women's empowerment (Samarakoon and Parinduri (2015); Cannonier and Mocan (2018); Kaffenberger et al. (2018)).

We implement a randomized controlled trial in 100 schools in cooperation with World Bicycle Relief (WBR). WBR provides bicycles to adolescent girls in grades 5, 6, and 7 who live more than 2.5 km away from school. The bicycle is given on the condition that it will be used primarily for attending school.[6] We randomly divided the schools into two treatment groups and a control group. The two treatment arms differ in the obligation of the parents/guardians to pay a small upfront payment at the time of receiving the bicycle. In the 'Payment Arm' (25 schools), the parent pays a small upfront amount ($\approx$ \$5)[7], while in the 'No Payment Arm' (20 schools) parents pays nothing (\$0). The remaining 55 schools are in the control group.

Optimal pricing of goods with large spillovers is of considerable policy interest (Morgan (2010)). Prices affect not just the demand for goods, but also their utilization. Proponents of the "cost-sharing" method argue that usage intensity will be higher as charging positive prices can have a psychological effect on consumers, inducing them to commit the "sunk cost fallacy" (Arkes and Blumer (1985);

[3]Zambia is ranked 116 out of 145 in the World Economic Forum's Global Gender Gap Index for 2015.

[4]Muralidharan and Prakash (2017) study a state-wide cycling program in the Indian states of Bihar and find that the policy increased girls' enrollment in secondary schools by 32% and reduced the gender gap by 40%. The study also finds an 18% increase in the number of girls who appear for the high-stakes secondary school certificate exam and a 12% increase in the number of girls who pass the exam.

[5]Borker (2018) finds that women in Delhi, India choose worse education outcomes for themselves in order to avoid street harassment, and that they are willing to incur an additional expense of USD 310 per year to travel by a route that in one SD safer.

[6]This is enforced by a Bicycle Supervisory Committee (BSC), which includes members of the community and the school, using their own set of rules.

[7]This is about 6.5% of the average rural monthly household expenditure in Southern Province (CSO (1991-2017)).

Thaler (1980)).[8] Thus, charging a small upfront cost in the 'Payment Arm' could induce parents to push their girls to use the bicycle more. On the other hand, charging a "zero price" can reduce improper usage of the product by invoking norms of social exchange (Ariely et al. (2018)). This can reduce occurrences of misuse of bicycles by the parents, making the ownership of bicycles more salient for the girls in the 'No Payment Arm'. Therefore, conceptually it is not obvious if the impact of the program will be larger for the 'Payment Arm' or the 'No Payment Arm'.

We measure the impact of the intervention on three first stage outcomes. First, girls in the pooled treatment group were 88% more likely to have access to a bicycle vis-à-vis girls in the control group. Second, we find that the intervention reduced the number of days the girls arrived late to school by 1.45 days in the previous week measured at the endline, which translates to a 66% reduction vis-à-vis the girls in the control group. Third, we find that the intervention reduced the daily distance cost of attending school. The time taken to commute to school decreased by about 35 minutes one way in the previous week measured at the endline, which translates to a 33% decrease from the baseline commuting time.

The intervention improved educational outcomes for girls in the treatment schools. We find that the intervention reduced overall absenteeism in the previous week by 28%, which translates to an addition of about 5 school days per academic year. These effects last for at least two years after the intervention.[9] In addition, we also find impacts on dropout, and grade transition two years after the intervention. Further, we find that the intervention increased the score for girls on a mathematics test by 0.11 *s.d*, but we find no impact on English test scores.[10] The effect size we find is consistent with the conditional and unconditional cash transfer literature, which concludes that the effects of these interventions on student achievement are small at best (Baird et al. (2013)).[11] We also find a small but positive effect on medium-run grade transition.

We find that the provision of bicycles led to an increase in empowerment outcomes for the girls in our treated schools. The intervention improved index of locus of control, bargaining, pro-sociality, and self-image, but it did not have a significant impact on the index of mobility, aspirations, or desired

[8]There are also additional selection effects of charging positive prices, which can increase usage intensity by screening individuals with the greatest propensity to consume (Oster (1995); Ashraf et al. (2010)), but can also dampen demand and reduce program coverage substantially (Cohen et al. (2010)). These effects are not applicable in our context since take up is 100% in all treatment arms.

[9]Students in primary school miss an average of 18 days in a school year in the Southern Province (DHS (2002)).

[10]Although moving test scores is non-trivial in the education literature, the strong association of the intervention with Mathematics test scores is consistent with the literature that finds Mathematics achievement to be more responsive to interventions changing curriculum or instruction time (see Cronin et al. (2005)).

[11]The meta-analysis by Baird et al. (2013) suggests a pooled effect sizes in the range of 0.04–0.08 standard deviations, respectively, for Unconditional Cash Transfer and Conditional Cash Transfer interventions.

fertility. In particular, we find that the index of locus of control increased by 0.16 *s.d* for girls in the treatment schools for the pooled treatment (Payment and No Payment combined). However, the intervention improved aspirations by 0.11 *s.d* and desired fertility by 0.18 *s.d* only for girls in the Payment Arm but not in the No Payment Arm.

Another important impact of the cycle intervention was on girls' safety. The intervention improved the perception of safety for the girls in the treated schools by 0.10 *s.d*. More importantly, measures of actual safety improved. Approximately 35% of the girls in baseline reported being teased on their way to school. We find that the intervention reduced the probability of girls being teased or whistled at on the way to school by about 22%, and the probability that a girl missed school or left for home early from school due to concerns of safety by about 33%.[12]

The translation of improvement in first-stage outcomes to an impact on *educational attainment* and *empowerment* outcomes crucially depends on the household response to bicycle provision.[13] In particular, the impact of the bicycle depends crucially on how girls are able to spend the time they save from commuting to school. We find that the intervention leads to a decrease in time spent on income-generating activities and school chores (for e.g. cleaning) for the girls in the treatment schools.

Finally, we also examine the heterogeneity in the impact of the program by baseline time taken to travel to school. Results suggest that the intervention relaxed the distance constraints for girls living further away from the school (as measured by the presence in the middle and top tercile), who experience a greater reduction in time taken to travel to school, and in the number of days they arrive late to school. However, it is the girls in the middle tercile who experience the greatest reduction in absenteeism compared to the bottom tercile. It is possible that in spite of the access to bicycles, the distance costs are still large for girls living the furthest from the school or that these girls face additional constraints.

This paper contributes to several related literatures, including to the literature on school access for female students. A standard policy response to address the problem of school access has been to construct schools. Several studies have shown the positive impact of school construction programs on enrollment and completion (Birdsall et al. (1985); DeTray (1988); Lillard and Willis (1994); Lavy (1996); Duflo (2001); Burde and Linden (2013); Kazianga et al. (2013); Azam and Saing (2017); Khanna (2019)). Our findings suggest that policies such as providing bicycles that directly improve school

[12]The results on safety are particularly important since any kind of violence negatively impacts access to education and a safe environment for learning and is part of the 1993 UN Declaration on the Elimination of Violence against Women.

[13]Das et al. (2013) find that household inputs are substitutes to school inputs, and that households reduce their spending to offset anticipated grants in the case of Zambia and India.

access by reducing the distance, and time cost of attending school, can be a viable short-to-medium run solution, especially when the trade-off between school access and scale is of first-order concern.

We also contribute to the large number of papers on the impact of conditional transfers on schooling outcomes. The majority of the well-identified studies on conditional transfer programs focus on cash transfers and find a positive impact on girls' enrollment and attainment (Fiszbein and Schady (2009); Barrera-Osorio et al. (2011); Baird et al. (2011); Filmer and Schady (2011); Heath and Mobarak (2015); Chaudhury and Parajuli (2010)), however, the evidence on test scores is weak (Baird et al. (2013)).[14]

Our findings also complement a limited number of experimental papers that directly aim at improving women's empowerment, however that evidence is mixed. Some of these interventions have looked at the impact of relaxing women's human capital constraints (Bandiera et al. (2018)), effect of role models (Beaman et al. (2009); La Ferrara et al. (2012); Riley (2017)), or strengthening women's financial control (Field et al. (2019)) on measures of economic empowerment, fertility, education, and labor force participation. Although the above-mentioned papers directly attempt to change women's empowerment, ours does this in a less-salient way through an improvement in education, aspiration, locus of control, desired fertility choices, and bargaining.

Finally, this paper also contributes to the growing debate on external validity around experimental and non-experimental studies (Dehejia et al. (2019); Gechter (2015); Vivalt (2019); Kowalski (2019)). The results we obtain here complement well with findings from Muralidharan and Prakash (2017), which was a non-experimental study of a large-scale cycling program for adolescent girls on enrollment and learning outcomes in India. The similarity of results suggests that identifying the underlying mechanisms can play an important role in understanding the challenges around replicability and external validity in international development.

The remainder of this paper proceeds as follows. In Section 2 we discuss the rural Zambian context. In Section 3 we present the design and methods we employ in this study. In Section 4 we discuss the data collection process and validity of the design. Section 5 details our empirical strategy and in Section 6 we present our results. Section 7 concludes with a discussion of the ways forward.

[14] See Fiszbein and Schady (2009); Rawlings and Rubio (2005) for a review of this literature, and Baird et al. (2013) for a review on relative effectiveness of conditional and unconditional cash transfers for schooling outcomes in developing countries.

2 Context

Zambia is a landlocked country and home to over 17 million people. The national education system in Zambia is divided into the following levels of education: primary education, as defined by the country, begins at age 7 and has a duration of 7 years (divided into lower basic grades (1-4) and middle basic grades (5-7). The entry age of lower secondary education (upper basic education) is 14 years, and it lasts 2 years. Upper secondary education begins at 16 years old and has a duration of 3 years. Tertiary or post-secondary education begins at age 19.

Zambia has made remarkable progress in improving access and equity in education, and provides close to universal education at the primary level, with the gross enrollment ratio (GER) of 108% at the primary level in 2013 (UNESCO (2016)). This puts Zambia's primary level GER above average for sub-Saharan Africa. Yet, despite this progress, there remain several challenges in Zambia, especially at the above primary level and for girls in rural areas: absenteeism, dropout, and performance (Mwanza (2015)). These disparities are further exacerbated between boys and girls throughout adolescence. For example, Zambia's GER for girls drops to 61% in the lower secondary level (Bank (2015)) and, starting in grades 6, significantly more girls leave school than boys (Bank (2015)), and students performance is among the lowest in the region according to the 2007 Southern Africa Consortium for Monitoring Educational Quality (SACMEQ).

Overall, despite striking progress in increasing overall enrollment, Zambia needs to reduce gender gaps in absenteeism, dropout, and educational attainment, for adolescents, especially at the above primary level. This will require addressing both the supply of and demand for education constraints. In this study, we partnered with World Bicycle Relief (WBR) and Ministry of General Education in Zambia to test the impact of a program that provided a bicycle to a school-going girl on measures of education and empowerment. This intervention mimics a conditional in-kind transfer program and had features of both demand and supply-side interventions (Muralidharan and Prakash (2017)). WBR provided bicycles if a student was enrolled, which is similar to demand-side conditional cash transfer programs, but access to bicycles reduces the daily distance cost and improves the safety of girls while going to school, which are similar to supply-side interventions.

3 Design and Methods

3.1 Treatment Arms

This experiment is a multi-treatment design with 100 schools randomly allocated to one of two treatment groups or a control group.

3.1.1 Payment Arm (T1)

The first treatment uses the same model as the Bicycle Education and Empowerment Program (BEEP) intervention that World Bicycle Relief (WBR) has rolled out in 19 districts in Zambia since 2009.[15] Students received a bicycle on the condition that the bicycle is used primarily to travel to school. A Bicycle Supervisory Committee (BSC) is in charge of monitoring the program at the school. A field mechanic was trained for each school, who provided maintenance checks and repairs for a fee borne by the recipient of the bicycle. Each school was then provided with a startup spare parts kit and the family of each beneficiary student was required to pay a contribution of 50 Kwacha ($\approx$ \$5) toward this kit. We randomly selected 25 schools for the Payment Arm.

3.1.2 No Payment Arm (T2)

The second treatment is a slight modification of the BEEP intervention. Students still received a bicycle on the condition that the bicycle be used primarily to travel to school, a BSC was formed, and a field mechanic was trained for each school. Each school was provided with a startup spare parts kit, but no contribution was obtained from the beneficiary students' families. We randomly selected 20 schools for the No Payment Arm.

We are thus removing only one part of the first treatment arm here: the upfront cost. The motivation behind this treatment arm comes from our focus group meetings and the psychology literature. A key take-away from the focus group meetings was that the ownership of the bicycle is more likely to be salient for a girl when parents do not have to pay any upfront cost. Ariely et al. (2018) suggests that charging a "zero price" can reduce improper usage of the product by invoking norms of social exchange. This is likely to reduce occurrences of misuse of bicycles by the parents, making the ownership of bicycles more salient for the girls in the 'No Payment Arm'.

[15] World Bicycle Relief, an organization that has worked in Zambia since 2009, and distributed over 183 thousand bicycles worldwide.

3.1.3 Control Group

Students in the control group did not received a bicycles. We randomly selected 55 schools for the control group.

3.2 Sample Selection and Randomization

World Bicycle Relief conducted an initial needs assessment in several districts in Zambia to identify three districts where students walked long distances to school, and where the program was not already being implemented. The three districts were Monze, Mazabuka, and Kalomo. The province that these three districts belong to is characterized by low population density and rural settlements, which makes distance a barrier for accessing basic services.

Within the three districts, a total of 100 schools were randomly selected from all public schools that met the following criteria.[16] First, each school had at least 20 girls enrolled in grades 5, 6 and 7, who walk more than 3 kms to school. Second, schools are basic schools, i.e., their starting grade is 1 or lower and their last grade is beyond grade 7 (end of primary) up to grade 9 (last grade before secondary education). WBR worked with the schools to compile a list of students in grades 5, 6 and 7, who walked more than 3 kms to school. The teachers generally knew the distance students walk and were able to reference information on where girls lived. In addition, we also confirmed the distance indicators during the baseline survey. From the list of eligible girls, in each school we randomly selected 25 of them to be part of our survey, for a total sample of 2,471 girls. Finally, within each district, the schools were randomly assigned to one of the two treatment arms or control.[17] Randomization was stratified by district. We describe the sampling procedure, and field protocols in details in Appendix Section B and C, respectively. In schools assigned to receive the intervention, WBR worked with the schools to select a Bicycle Supervisory Committee (BSC), consisting of 10-12 members; teachers, Parents and Teachers Association (PTA) members, local leaders and student representatives. All girls belonging to the original list of eligible girls (not only the 25 that were part of the survey) received a bicycle.[18]

[16] Monze (44 schools out of a total of 135 schools), Mazabuka (20 schools out of 81), and Kalomo (36 schools out of 124).

[17] Appendix Figure A.1 shows the study districts and the selected schools for the sample.

[18] If there were multiple girls from the same household, only one bicycle was given to them.

3.3 Timeline

Baseline data was collected during the second term of the 2017 school year (July to August 2017).[19] The baseline was first launched in the district of Monze and our team spent 9 days visiting all of the 44 schools. The team then moved to Mazabuka and spent 4 days to visit all of the 20 schools in the district in 4 days. Finally, the team spent 8.5 days to visit all of the 36 schools in the district of Kalomo.

The girls were asked questions on school attendance, mobility, and measures of bargaining power. In addition, they were tested on basic skills in Mathematics and English, and a measure of focus/attention. School characteristics like enrollment and infrastructure were collected from the headteachers. After the collection of the baseline, schools were randomly assigned to one of the two treatment arms or control.

The bicycles were distributed to girls in schools during the third term of the 2018 school year (September to November 2018). Distribution was done with all girls at the same time in a special event organized by WBR.

The endline survey was implemented one school year after the bicycles were distributed. Similar to baseline, the girls were asked questions on school attendance, mobility, and measures of bargaining power, as well as tested on basic skills in Mathematics and English, and a measure of focus/attention. School characteristics like enrollment and infrastructure were collected from the head teachers during the endline. In a sub-sample of schools, we also collected non-survey measures of girls' bargaining power within the household and their willingness to share an opinion in a group.

See Figure A.2 for the timeline of the intervention.

4 Data Collection and Validity

The empirical analysis uses both survey and administrative data collected from schools, students, and head-teachers, over the course of the study (see Appendix Section C for Field Protocols). The research team used a variety of methods to collect data from students, and head-teachers (or acting head-teachers).

[19] It took place between 5^{th} July and 10^{th} August 2017, and the team (supervisors and surveyors) worked 21.5 days in the schools, over a period of 5 weeks.

4.1 Surveys

We administered a 40 minute face-to-face survey to girls in the sample using a tablet to collect data on primary outcomes (see Appendix Section A for list of variables collected), and a paper-based survey to collect data on sustained attention (D2 Test) which lasted 10 minutes, learning assessment in English and mathematics which lasted 25 minutes, and a semi self-administered survey to collect data on questions that were sensitive, which lasted 10 minutes. We also administered a 40 minute face-to-face survey to head-teachers (or acting head-teachers) using a tablet to collect data on school characteristics. Throughout the girls surveys, particular care was taken to ensure privacy: girls' were interviewed by themselves, without interference from teachers, head-teachers, or classmates. In particular, only after the face-to-face survey ended with a girl in the school, the enumerator started the next survey.

4.2 School Records Data

The research team collected administrative data on student attendance, enrollment, and dropout in 2019 (two years after the start of the intervention). The survey team took pictures of the attendance registers which were entered manually for the analysis, and also collected data on girls enrollment, dropout, and grade transition from the District Education Boards Secretaries (DEBS).

4.3 Primary Outcome Variables

We pre-specified the following primary outcomes for the endline (see Appendix Section A for list of primary outcomes) in the pe-analysis plan (PAP) registered at RCT ID: AEARCTR-0003339.

Educational Outcomes: We use three measures of educational outcomes. They are: *(i) Days Absent*, which is a self-reported data on days absent in the previous week *(ii) Dropout*, which is a dummy variable is the girl dropped from the school at the endline and *(iii) Grade Transition*, which is a dummy variable if the girl progressed to a higher grade, conditional on not dropping out.

Empowerment Outcomes: We use girls' responses to 23 individual questions (Appendix Section A lists the individual variables used in the indices) on various indicators of empowerment to construct the following 4 indices: *(i) Index of Mobility and Safety (ii) Index of Aspiration (iii) Index of Locus of Control* and *(iv) Index of Marriage and Fertility.*[20] These indices include variables in which the

[20]The research team used a self-administered survey for questions related to safety, security, marriage, and fertility to allow more privacy to the respondent and to avoid social desirability bias.

higher value indicates a better or positive outcome. We interpret a positive value in the index as higher empowerment. The variables used to measure empowerment of girls has been validated and used by Kabeer (1999); Laszlo and Grantham (2017); Dhar et al. (2018).[21]

Each index mentioned above is constructed by aggregating responses to several individual questions into an index, which is a weighted average value of the individual variables, with weights constructed by normalizing the variables to have the same $s.d$, and then recovering the weights from the inverse covariance matrix, following the approach described in Anderson (2008) (for details on the steps for index construction see Appendix Section E).

4.4 Secondary Outcome Variables

Similar to the primary outcomes, we pre-specified the following secondary outcomes for the endline (see Appendix Section A for list of primary outcomes) at RCT ID: AEARCTR-0003339.

First Stage Outcomes: We first report the impact of cycles on various measures of distance cost of attending school: *(i) Days late to school in the previous week (ii) Access to the cycle* and *(iii) Travel time to school*.

Behavioral Outcomes: We use girls' responses to 16 individual questions (Appendix Section A lists the individual variables used in the indices) on various indicators of behavior to construct the following 3 indices using Anderson (2008): *(i) Index of Bargaining (ii) Index of Pro-Sociality* and *(iii) Index of Pro-Sociality*.

Test Scores: We administer *(i) English* and *(ii) Mathematics* test to measure student learning outcomes. Both tests were based on the tests administered at the national level by the Examination Council of Zambia for Grade 5 (see Appendix Section F for the list of questions).

Index of Focus: We measure focus of girls, using D2 test. The D2 test consists of 14 lines with 47 characters in each line. This character is a letter "d" or "p" marked with small dashes either above or below. The respondent has 30 seconds per line to circle the letter "d" with two marks, above or below in any order (see Appendix Section G for an example of the test).

The D2 test measures *(i) Speed*, which is the total number of observations processed in the D2 test and *(ii) Accuracy*, which is the correct number of observations processed in the D2 test. We also combine these to create an *Index of Focus* using Anderson (2008).

[21] See Glennerster et al. (2018) for a practical guide to measuring women's and girls' empowerment in impact evaluations.

Time Use: We measure how girls spend their time on different activities during a normal weekday (Appendix Section A lists the individual variables to measure various activities). These activities are: *(i) School Chores (ii) Extra-curricular Activities (iii) Studying and Homework (iv) Household Chores (v) Engaging in Income Generating Activities* and *(vi) Spending time with Friends.*

These are categorical variables that specify the amount of time spent by the girl on a particular activity. They take the value of 0 if the girl spent no time doing that activity, 1 if she spent less than 30 minutes, 2 for between 30 and 60 minutes, 3 for between 60 and 90 minutes, 4 for between 90 and 120 minutes, and 5 for more than 120 minutes.

4.5 Integrity of the Experimental Design

4.5.1 Baseline Balance

We report the baseline characteristics of the schools by treatment status in Appendix Table A.1 and find that the sample is balanced across most variables, except the number of girls' toilets in the school. The first panel reports the mean and the standard deviations of the baseline variables for the schools in the respective treatment arms, and the second-panel tests for statistical difference across the two arms. The schools in our sample on average have an enrollment of 680 students, equally split between boys and girls. They employ an average of about 13 teachers, also equally split between male and female teachers, about three-quarters of who reside at the school premises. A large fraction of the schools ($\sim$ 60%) have a sanitation program running, which is also confirmed by the presence of separate toilets for girls and boys. Most schools have a computer lab ($\sim$ 95%). The schools are similar in the degree of "remoteness", as can be seen from the distance to closest town (47 km), tarmac road (25 km), and secondary school (20 km).

Similarly, we report the baseline characteristics of the girls in the two treatment arms in Appendix Table A.2 and find that the sample is balanced across most variables, except grade, number of times girls eat meat in the meal, socio-economic index, and self-esteem index. The first panel reports the mean and the standard deviations of the baseline variables for the girls in the respective treatment arms, and the second-panel tests for statistical difference across the two arms. The average age for girls in our sample is 13, with about 15% of them currently engaged or married, and 5% being ever pregnant. The girls have a family size of at least 6 members, and 80% of them have both parents alive. A third of the girls have repeated a grade in the past. These girls come from extremely impoverished

backgrounds, where they have not had enough food to eat in 1 out of the last 7 days. The girls are similar across the treatment arms in terms of the indices on the locus of control, self-esteem, and gender attitudes. These indices have been constructed using Principal Component Analysis. Appendix Section A has a description of the variables in each of these indices, and their spread across treatment arms.

We present a detailed description of the sampling of schools and girls in Appendix Section B. Overall, we conclude the randomization was successful.

4.5.2 Compliance

The allocation of the bicycles was carried out by World Bicycle Relief, in partnership with the Ministry of General Education in Zambia in the 45 schools in treatment group. The distribution took place within the schools and was considered successful by WBR and the research team. Only one girl out of 2,471 refused the cycle.

4.5.3 Attrition

Attrition on outcome measures is below 9%. In baseline, we interviewed 2,471 girls. In the endline, we used a two-stage tracking method. This method consists of a first step in which we attempt to interview the girls in the school where they were enrolled at baseline. We found 72% of the girls in this phase. From those not found in school, we select 50% randomly to track in their households, villages, and, if necessary, in other districts. We weight the answers from those girls interviewed in the second phase depending on the probability of being sampled in the analysis. This method allows us to maximize the resources available and keeping the effective attrition rate as low as possible. In the first line of Appendix Table A.3, we present the attrition rate if we do not take into consideration the weighting of the girls found in the second stage of tracking by treatment arm. In the last row, we show the effective rate of attrition, using the method described. The tracking rate of girls in the control group is lower (90%) than the one in the two treatment arms (94% and 92%).[22] We find significant differential attrition between the control group and the No Payment Arm. However, this difference is not a concern given that the tracking rate is above 90% for all the groups.

Nevertheless, we check if the attritors are different from those that we interviewed in the endline

[22]In Appendix Figure A.3 we present the numbers of girls that we interviewed in different phases of the project by treatment.

in any of the observable characteristics from baseline. In Appendix Table A.4, we see that none of the indices of observable at baseline (socioeconomic characteristics of the household, index of self-esteem, index of locus of control, and index of gender attitudes) are significantly different for those students interviewed at baseline and for those missing. Finally, we also estimate Lee bounds of the treatment effects using Lee (2009) to consider the possibility that other non-observable are endogenous to the treatments. We present the results in Appendix Tables A.5–A.7 and find minimal differences between upper and lower bounds of the treatment effects. Overall, due to the low amount of attrition, our main results are robust to Lee bounding.

4.5.4 Data Analysis

We evaluate the impact of providing a bicycle to a school-going girl on various measures of education and empowerment outcomes. We collected a large number of outcomes through a primary survey to study the impact of the intervention. For complete transparency, we follow the Pre-Analysis Plan (PAP) which is available and timestamped at RCT ID: AEARCTR-0003339.[23] The PAP specifies the variables to be analyzed, construction of indexes, how we plan to address multiple inferences, the empirical specifications we plan to use, and our approach for tracking and the handling of attrition. The empirical analysis reported in this paper follows the PAP. The paper also reports results where we deviated from the PAP. In such cases, we report these deviations and provide reasons for them while discussing those results.[24]

5 Empirical Strategy

5.1 Reduced-Form Specifications

We estimate the intent-to-treat (ITT) impact of the cycles program for both the treatments together (Payment Arm and No Payment Arm combined, relative to the control), and separately using the following specification:

$$Y_{i,s,t=1} = \alpha_0 + \alpha_1 T_s + \alpha_2 Y_{i,s,t=0} + X_{i,s}\theta + C_s\gamma + e_{i,s} \quad (1)$$

[23] PAP registered at AEA RCT Registry: https://www.socialscienceregistry.org/trials/3339/history/34596

[24] A recent paper by Banerjee et al. (2020) discusses the costs and benefits of adhering to PAP, and recommends that the final research paper be written and judged as a distinct object from the "results of the PAP".

where $Y_{i,s,t=1}$ is the outcome variable of student i in school s measured in post treatment, and $Y_{i,s,t=0}$ is its baseline value. T_s is an indicator for the school assigned to the treatment group, $X_{i,s}$ is a vector of individual level controls, C_s is a vector of school level controls, and $e_{i,s}$ is the error term. We cluster the standard errors at the school level in order to account for unobservable correlation in girls in the same school, and also because the treatment was assigned at the school level. For outcomes where we do not have baseline values, we do not include $Y_{i,s,t=0}$ in the estimation. Finally, β_1 is our main coefficient of interest and provides the ITT effect, which is the effect of being given a bicycle in 2017 on the outcome variable.

We further present the ITT impact of the cycles program by the two treatment arms, relative to the control using the following specification:

$$Y_{i,s,t=1} = \beta_0 + \beta_1 Payment\ Arm_s\ (T1) + \beta_2 No\ Payment\ Arm_s\ (T2) + \beta_3 Y_{i,s,t=0} + X_{i,s}\theta + C_s\gamma + e_{i,s} \quad (2)$$

where $Payment\ Arm_s\ (T1)$ and $No\ Payment\ Arm_s\ (T2)$ are indicators for being assigned to each of the treatment groups and all other variables are the sample as in Equation 1. Finally, β_1 and β_2 will provide the ITT effects for each of the two treatment groups.

5.2 Accounting for Multiple Comparisons

This study entails a large number of outcome variables, therefore it is important to address the concern of false positives. We address this using two methods.

First, to control the probability of Type I errors, we use the Benjamini and Hochberg (1995) False Discovery Rate correction (as specified in the PAP).[25] We show results with the corrected p-values, where the corrections are made within outcome groupings. So, the primary outcomes will all be corrected together, secondary outcomes corrected together, and heterogeneity tests will be corrected together. Second, we compute standardized indexes using Anderson (2008) for several primary and secondary outcome variables, as specified in the PAP. We *do not* apply this correction and use the conventional statistical significance levels while discussing individual variables within particular outcome variables.

[25] In addition, we also use Westfall et al. (1993) and Holm (1979) to address the concern of false positives (not specified in the PAP).

5.3 Minimum Detectable Effect Sizes

We present the minimum detectible effect sizes (MDEs) in Appendix Table A.8 for our primary outcome variables. Due to very low intra-cluster correlation, the MDEs are generally very low. We are powered to observe changes at or well below 20% for all outcomes.

6 Results

6.1 Impact of Cycles on First Stage Outcomes

We first report the impact of cycles on first stage outcomes: *(i) Days late to school in the previous week (ii) Access to the cycle* and *(iii) Travel time to school*, in Table 1.[26]

Panel A in Table 1 reports the results for the pooled treatment estimated using Equation 1. We find that the intervention reduced the number of days the girls arrived late to school by 1.45 days (Column 1) in the past week, which translates to a 66% reduction vis-à-vis the girls in the control group, and is a direct result of the decline in the average commute time to school. Most importantly, we find that the girls in the treatment group were 88% (Column 2) more likely to have access to a bicycle vis-à-vis girls in the control group. Finally, the intervention reduced the commute time to school by 35 minutes (Column 3) one way. This corresponds to a decrease of one-third of the control group (baseline) commuting time. Overall, it is reassuring to find that the intervention had large and statistically significant impact first stages outcomes.[27]

Panel B reports the impact of the intervention by the two treatment groups separately using Equation 2. We do not find any statistical difference in the outcome measures between the two treatment groups. This is not surprising since the impact of cycles on first stage outcomes should not vary by Payment and the No Payment treatment arms. All estimated impact remains significant when the p-values are corrected for multiple hypotheses testing.

[26] Table A.9 provides the means and standard deviations of the variables in the estimation sample.

[27] Primarily motivated by the likely decrease in time taken to reach school due to cycles, we present the impact of the intervention on an index of focus in Table A.10. This index measures *(i) Speed* and *(ii) Accuracy* using a D2 test. We do not find a statistically significant impact on index of focus, however, the signs are positive.

6.2 Impact of Cycles on Educational Outcomes

Table 2 reports the impact of the intervention on educational outcomes pre-specified in the PAP.[28] Panel A presents the results for the pooled treatment and Panel B reports the same separately by the two treatment groups.

We find that the cycles intervention reduced overall absenteeism in the previous week (Column 1 in Panel A) by 28%, with the effect sizes being identical for the two treatment arms (Panel B). The impact is sizeable, as this translates to an addition of about 5 additional school days for girls in the Southern Province, where students in primary school miss an average of 18 days in a school year (DHS (2002)). We do not find any impact on dropouts (Column 2) and grade transition (Column 3), although the sign of the coefficient is negative and in the right direction. The result on dropouts is not surprising given the low level of dropout in the study sample (the control group mean is 6%). Similarly, grade transition is automatic up to grade 7 in our sample schools. Although statistically insignificant, taken together, the results suggest that girls in treatment are more likely to be enrolled in school, but they do not necessarily graduate to higher grades. Given that grade transition is automatic up to grade 7, this implies that the intervention is successful in keeping those girls in school who would have otherwise dropped out during the transition to secondary school from grade 7 to grade 8.

Given the sizable reduction on days late to school in the previous week (Column 1 in Table 1) and days absent (Column 1 in Table 2), it is imperative to look at the impact on test scores. We administered tests in English and Mathematics to girls in the baseline and the endline.[29] We find that the intervention increased the Mathematics score for the girls by 0.11 *s.d* (Column 4), which is statistically significant. However, we do not find any impact on English test scores (Column 5). These results are in line with the literature that finds Mathematics achievement to be more responsive to interventions changing curriculum or instruction than English (Cronin et al. (2005)). Results in Panel B suggests that the improvements in learning outcomes are driven by girls in the Payment Arm, although there is not enough evidence to support the differential treatment effect between the two treatment arms.

Overall, the impact on Mathematics tests score is noteworthy. First, theoretically, it is not obvious if the reduction in days late to school and days absent (mechanically both increase the instruction time)

[28] We pre-specified *Days Absent*, *Dropout* and *Grade Transition* are registered as primary outcomes; and test score on *English* and *Mathematics* as secondary outcomes in the PAP registered at RCT ID: AEARCTR-0003339.

[29] The English and Mathematics tests were based on tests administered at the national level by the Examination Council of Zambia for grade 5.

will necessarily improve test scores, as it depends on how this enters the students' education production function.[30] This is particularly challenging at the post-primary level. In fact, the review paper by Glewwe and Muralidharan (2016) suggests that many expensive standard school inputs are often not effective at improving learning outcomes, compared to the interventions that focus on improved pedagogy and school governance. Second, the effect size we find is consistent with the conditional and unconditional cash transfer literature, which conclude that the effects of these interventions on student achievement are small at best (Baird et al. (2013)).[31]

6.3 Medium Run Impact of Cycles on Educational Outcomes

To study the medium-run impact of the intervention on girls' education, we collected administrative data on students' attendance, enrollment, dropout, and grade transition two years after the cycle intervention.

Table 3 reports the impact of the intervention on girls' attendance in the term I and term II of 2019. Panel A presents the results for the pooled treatment and Panel B reports the same separately by the two treatment groups. We find that the intervention continued to improve girls' attendance two years after the cycles program. In particular, girls in the treatment schools miss 3.27 days less (39% reduction) in term I, and 3.13 (37% reduction) days less in term II, compared to the girls in the control school in Panel A of Table 3. We do not find the impact to vary by the two treatment groups in Panel B of Table 3.

Table 4 reports the impact on girls enrollment, dropout and grade transition (enrollment and dropout are the same). Panel A presents the results for the pooled treatment and Panel B reports the same separately by the two treatment groups. We find that the girls in the treatment schools are 11% more likely to be enrollment (or 18% less likely to be dropped out) than the girls in the control schools. More importantly, the intervention increased the rate of grade progression of girls in the treatment schools by almost 3% compared to the control schools. Once again, we do not find the impact to vary by the two treatment groups in Panel B of Table 4.

From a policy standpoint, these are very encouraging results since the intervention continued to improve girls' educational outcomes in medium run.

[30] Non-experimental evidence from developing countries suggests weak evidence of an increase in additional days of instruction on test scores (Aguero and Beleche (2013); Bellei (2009)).

[31] The meta-analysis by Baird et al. (2013) suggests a pooled effect sizes in the range of 0.04–0.08 *s.d*, respectively, for Unconditional Cash Transfer and Conditional Cash Transfer interventions.

6.4 Did Bicycles Transform Girls' Lives?

6.4.1 Impact on Measures of Empowerment

Tables 5 and 6 reports the impact of the intervention on indices of empowerment pre-specified in the PAP. *Index of Mobility and Safety*, *Index of Aspiration*, *Index of Locus of Control* and *Index of Marriage and Fertility* are registered as primary outcomes, and *Index of Bargaining*, *Index of Pro-Sociality* and *Index of Pro-Sociality* as secondary outcomes. Panel A presents the results for the pooled treatment and Panel B reports the same separately by the two treatment groups.

Overall, estimates from the pooled treatments (Panel A in Tables 5 and 6) suggests that the intervention did not have any impact on the indices of: mobility and safety, aspiration, and marriage and fertility. However, the intervention did improve the indices of: locus of control, bargaining, pro-sociality, and self-image. In particular, the index of: locus of control improved by 0.16 *s.d* (Column 3 in Table 5), bargaining improved by 0.19 *s.d* (Column 1 in Table 6), pro-sociality improved by 0.14 *s.d* (Column 2 in Table 6), and self-image improved by 0.11 *s.d* (Column 3 in Table 6) for girls in the treatment schools.

These are important results as the index of locus of control measures the degree of control the girls believe they have over outcomes in their lives, and how satisfied they are with their life in general. Similarly, the index of bargaining measures whether girls have access to and control over small amounts of resources, the clothes they wear, the food they eat, and whether they are able to discuss matters pertaining to their lives with their parents.[32] The index of pro-sociality measures the participation of girls in local clubs, their willingness to help out their friends, and their knowledge of the local leadership, while the index of self-image measures what girls think of themselves vis-à-vis their peers in term of academic achievement, and their probability to succeed in future.

Panel B in Tables 5 and 6 reports the impact by Payment Arm and No Payment Arm separately. We find that the intervention improved the index of aspiration and fertility and marriage for girls in the Payment Arm. In particular, we find that the cycles intervention improved the index of aspiration by 0.11 *s.d*, and index of fertility and marriage by 0.18 *s.d* (Columns 2–3 in Panel B of Table 5). The index of aspiration measures the girls' aspirations with regard to the years of education they want to receive, and their future participation in the workforce. Similarly, the index of fertility and marriage

[32]This index is slightly different from the one we registered in the PAP. In particular, we excluded two variables that measure rebellion. The estimated coefficient for the pooled treatment was 0.10 (0.06) and statistically significant at 10 percent, for Payment Arm was 0.07 (0.07) but not statistically significant, and for No Payment Arm was 0.14 (0.07) and statistically significant at 5 percent.

measures the girls' desired fertility behavior, and preferences on the age of marriage.

Improvements in the index of aspiration, and fertility and marriage suggest that the girls in the Payment Arm aspire to continue education and participate in the workforce in the future, and have a lower preference to get married early and having many children. In our context, when a credit-constrained household spends money ex-ante to acquire the bicycles for their girl child, this might be a signal for the girls in the Payment Arm that their parents are invested in their education and general well-being. Furthermore, charging a small upfront cost in the Payment Arm is likely to induce parents to push their girls to use the bicycle more. This is consistent with the "sunk cost fallacy" which argues that usage intensity is higher when consumers are charged positive prices (Arkes and Blumer (1985); Thaler (1980)).

Taken together, these results provide the *first causal evidence* on the transformative role a bicycle can play on women empowerment, which was first highlighted by scholars of the women's suffrage movement.[33] The positive impact of the intervention on the empowerment of girls' in rural Zambia is likely to have a lasting impact on their future living standards as argued by Duflo (2012). More importantly, the results on measures of empowerment directly contribute to the policy debate on interventions that can improve female empowerment, which is a priority for policymakers in developing countries (UN Sustainable Development Goal No. 5).[34]

6.4.2 Impact on Safety

Table 7 reports the impact of the intervention on measures of safety. We disaggregate the index of mobility and safety (Column 1 in Table 5) into sub-indices of mobility and safety separately. These were not pre-specified in the PAP. Panel A presents the results for the pooled treatment and Panel B reports the same separately by the two treatment groups.

The index of mobility and safety estimated in Column 1 of Table 5 contains questions related to whether (and the intensity) the girl is allowed to visit friends, family members or going to the market by herself; and questions related to whether the girl feels safe moving around in the village

[33] In 1895, at the age of 80, suffragist leader Elizabeth Cady Stanton claimed that "the bicycle will inspire women with more courage, self-respect, self-reliance..." Stanton predicted the power of the bicycle in transforming the lives of women, realizing that the independence women were gaining because of this invention would allow for growth in other areas of their character. Having the ability to be fully self-reliant, often for the first time in their lives, would encourage women to be more courageous in other areas, such as demanding voting rights.

[34] The Gender Parity Goals of the UN Sustainable Goal No. 4 aims to achieve gender equality and empower all women and girls. Similarly, Goal No. 4 aims to ensure inclusive and equitable quality education and promote life-long learning opportunities for all.

and traveling to school. Using these two sets of questions, we create sub-indices of mobility and perceived safety to separately report the impact of the intervention in Table 7 (Columns 1–2). We find that the intervention did not have an impact on the sub-index of mobility (Column 1), but improved the perceived safety of girls in the treatment schools by 0.10 s.d. Once again, we do not find any statistical difference in the outcome measures between the Payment and No Payment Arm. Although it is surprising that the intervention did not improve the mobility of girls in the treatment schools, upon further investigation, we found out that parents of the girls in the treatment schools as well as the members of the Bicycle Supervisory Committee (BSC) considered the bicycle to be a precious asset, to be used by the girls only for the purpose of traveling to school. These restrictions on how the bicycle should be used by the girls explain why we do not find an impact on girls' mobility.

In this study, we also asked questions on two measures of actual safety.[35] We find that the intervention reduced the probability that a girl is teased or whistled at on the way to school by 22% (Column 3). More importantly, we find that the intervention reduced the probability that a girl missed school or left for home early from school due to concerns of safety by 33% (Column 4). Reassuringly, we do not find these results (Columns 3–4 in Panel B) to differ by the treatment arm.

The results on girls' safety are particularly important as approximately 35% of the girls report having been sexually harassed on their way to school in the baseline. There is strong evidence that violence against women, including sexual harassment, has a negative impact on psychological costs (Langton and Truman (2014)), human capital investments (Borker (2018)), and labor force participation (Siddique (2018)). Also, harassment by strangers strongly impacts women's perception of safety across social contexts (Macmillan et al. (2000)). We believe that the improvements in safety due to bicycles is likely to have a far-reaching impact on girls' well-being.

6.5 Impact of Cycles on Time Use

Table 8 reports the impact of the intervention on measures of time-use pre-specified in the PAP. *School Chores*, *Extra-curricular Activities*, *Studying and Homework*, *Household Chores*, *Engaging in Income Generating Activities*, and *Spending time with Friends* are registered as secondary outcomes. Panel A presents the results for the pooled treatment and Panel B reports the same separately by the two treatment groups.

In Table 8 we estimate the impact of the intervention on changes in the way the girls spend their

[35] These two outcomes are not part of the index of mobility and safety and were not pre-specified in the PAP.

time on different activities during a normal weekday. The dependent variables are categorical variable, specifying the amount of time spent by the girl on a particular activity.[36] The estimated coefficients are odds ratio from an ordered logit specification. We do not find any impact of the intervention on most of these measures of time use, except time spent in Engaging in Income Generating Activities (Column 5). The odds of spending time in Income Generating Activities for the girls in treatment schools is 0.75 times that of girls in the control group, i.e., girls in treatment are less likely to be engaged in income generating activities.

A plausible explanation could be that the intervention had positive income effects for the family, thereby not requiring the girl to engage in any income generating activity, or the intervention changed the value that the family places on education vis-à-vis engaging in short-term income generating activities, which resulted in a decline in the amount of time spent by the girls in these activities. However, we do not see an increase in the time spent by girls in studying, but this could be because we measure time in very coarse intervals.

6.6 Heterogeneity Analysis

Tables 9 and 10 reports the heterogeneous impact of the intervention by the baseline time taken to travel to school (we interact tercile of the baseline time taken to school by the treatment groups) as pre-specified in the PAP.

We do not find any heterogeneous impact of the intervention on access to cycles (Column 1 of Table 9). This is consistent with the theory of change since it implies that after a year, girls living closer or further away from the school are equally likely to still have access to the cycle. Girls in the middle tercile experience greater reduction in absenteeism vis-à-vis the bottom tercile (Column 2 of Table 9), however there is no statistically different impact for girls in the top tercile. As expected, girls living further away from the school (as measured by the middle and top tercile), experience greater reduction in time taken to travel to school (Column 3 of Table 9). Similarly, girls in the middle and top tercile also experience greater reduction in the number of days they arrive late to school (Column 4 of Table 9). We cannot reject the null of no impact on absenteeism for the girls in the top tercile (p-value = 0.59). Finally, we do not find statistically different heterogeneous impact on learning outcomes (Columns 5–6 of Table 9), but the direction of the impact is consistent with our results on school

[36]The dependent variable takes the value 0 if the girl spent no time doing that activity, 1 if she spent less than 30 minutes, 2 for between 30 and 60 minutes, 3 for between 60 and 90 minutes, 4 for between 90 and 120 minutes, and 5 for more than 120 minutes.

absenteeism.

We further estimate the heterogeneous impact of the intervention by the baseline time taken to travel to school on measures of empowerment. We find that this intervention improved the *index of locus of control* (Column 3 of Table 10) and *bargaining* (Column 5 of Table 10) for girls living in the bottom and middle tercile. However, we do not find an impact on girls in the top tercile.

Overall, the intervention relaxed the distance constraints for the girls living further away from the schools the most, which translates to improvements on the intensive margin but does not necessarily translate to changes in the extensive margin for girls in the top tercile. It seems like that for girls living furthest from the school the distance costs are still binding, or they face additional constraints, and a possible long-term solution would be school construction. Taken together, it seems like the improvement in measures of empowerment is routed through increased attendance in school and not through the access and ownership of the cycle.

7 Conclusion

In this paper, we find that a conditional kind transfer of a bicycle can be a useful policy tool to transform the lives of girls: *girls in the treatment schools reported less commute time to school, absenteeism, late arrival to school, and improved test scores and safety, time use, grade transition, and dropout.*

Scholars of the women's suffrage movement, including Susan B. Anthony in the United States, have highlighted the role played by bicycles in empowering women in the 19^{th} century. Consistent with this historical perspective, we find that the provision of bicycles improved girls' empowerment through *improved locus of control, bargaining, pro-sociality, and self-image.* Policies that aim at improving female empowerment have limited success due to the deeply rooted cultural norms that lead to discrimination against women in all spheres of life (Jayachandran (2015)). Work by Duflo (2012) suggests that improving female empowerment may also have a lasting impact on women's future living standards. It is worth highlighting that finding positive impacts on both educational and empowerment outcomes is perhaps the most unique result of this intervention.

In rural Zambia, approximately 35% of the girls report having been teased on their way to school (Fiala et al. (2020)). We find that the intervention improved the safety of girls in the treatment schools. In particular, it reduced the probability of whether girls were teased or whistled at on the way to school by about 22% and reduced the probability that a girl missed school or left for home early from school

due to concerns of safety by about 33%. Given the negative impact of sexual harassment on women's access to education and learning (Borker (2018), improvement in safety due to bicycles is likely to have a far-reaching impact on girls' well-being.

We further find that the intervention led to a decrease in time spent on income-generating activities for the girls in the treatment schools. A plausible reason could be that the intervention had positive income effects for the family, thereby not requiring the girl to engage in any income-generating activity, or the intervention changed the value that the family places on education vis-à-vis engaging in short-term income-generating activities. However, we do not find a change in any other activity, including time spent in performing household chores or studying, which could be because we only measure time in very coarse intervals.

Results from the two treatments arms – Payment vs. No Payment – suggest that girls who received bicycles with the small cost to their family had higher levels of aspirations, self-image, and a desire to delay marriage and pregnancy. Surprisingly, we do not find these effects in the zero cost treatment. We believe these results are due to girls perceiving the payment from the family as a desire to increase future investment in them. Although we do not have measures of how the parents feel about investment in the girls, based on discussions with parents we do not believe that these expectations from the girls are likely correct. We believe that future work is needed on the disconnect between parents and child education expectations.

While an intervention like providing bicycles to students is likely too expensive for most governments, and there are more cost-effective ways to increase schooling for girls, such as paying school fees, we believe our results have several important policy implications. First, we show that a policy like the provision of bicycles that improves access to school through a reduction in distance costs can improve educational outcomes, at least in the short run. This is especially important since discriminatory social norms that limit girls' access to education and labor force participation remain a challenge in many parts of the world. In addition, school construction programs, a default approach to address the access to school problems, is expensive, takes a long time to complete, and might not be cost effective. Second, we demonstrate that the intervention improved girls' empowerment, which directly contributes to the "Gender Parity" objective of the UN Sustainable Development Goals 2030. Third, we did not find any differential impact between the two treatment arms that is statistically significant. Many policies are designed with "conditionality", which is administratively burdensome and increases the overall cost, especially from both implementation and monitoring points of view. In our context, we

did not find evidence that a small upfront payment had any "unintended" impact on girls' outcomes. Finally, we contribute to existing research on conditional cash transfers, which are increasingly used as a policy tool to increase female schooling in developing countries, and find that a bicycle improves girls' educational outcomes.

Taken together, the results from Muralidharan and Prakash (2017), a non-experimental evaluation of a large-scale cycling program for adolescent girls in India, and this paper, a randomized control trial in Zambia, point to a potential policy tool that can address both the gender gap in education and improve female empowerment. More broadly, results from this paper suggest that identifying the underlying mechanisms can play an important role in understanding the challenges around replicability and external validity in international development. This is especially important since results of randomized evaluations of the same intervention vary substantially across trials (Vivalt (2019)), and even within the same location, causal impact varies due to random variation in conditions over time (Rosenzweig and Udry (2020)).

Future research could focus on studying the long-term impact of such policies on girls' age of marriage, fertility decisions, bargaining, and the community level spillovers on norms and aspirations from an in-kind transfer that went to a population that does not normally receive items of relatively high value. More broadly, it is important to study how policies aimed at improving girls' education impact community dynamics and norms regarding girls' education. This is especially important since communities can have their own norms regarding girls' education. In the context of Sub-Saharan Africa, apart from problems of access, income, and information, girls face additional cultural constraints like early marriage and pregnancy. Such social expectations and gender bias can lead to certain family practices that deprive girls of not just educational opportunities, but also alter their aspirations. Finding innovative ways of changing social norms might prove to be a sustainable way of tackling the problem of higher dropouts for adolescent girls in developing countries. Though social norms are slow-moving and hard to change, previous research has shown evidence of change in norms in response to exposure (Dhar et al. (2018)). ■

References

Aguero, J. and T. Beleche (2013): "Test-Mex: Estimating the Effects of School Year Length on Student Performance in Mexico," *Journal of Development Economics*, 103, 353–361.

Anderson, M. L. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103, 1481–1495.

Ariely, D., U. Gneezy, and E. Haruvy (2018): "Social Norms and the Price of Zero," *Journal of Consumer Psychology*, 28, 180–191.

Arkes, H. R. and C. Blumer (1985): "The Psychology of Sunk Cost," *Organizational Behavior and Human Decision Processes*, 35, 124–140.

Ashraf, N., J. Berry, and J. M. Shapiro (2010): "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia," *American Economic Review*, 100, 2383–2413.

Azam, M. and C. H. Saing (2017): "Assessing the Impact of District Primary Education Program in India," *Review of Development Economics*, 21, 1113–1131.

Baird, S., F. H. Ferreira, B. Özler, and M. Woolcock (2013): "Relative Effectiveness of Conditional and Unconditional Cash Transfers for Schooling Outcomes in Developing Countries: A Systematic Review," *Campbell systematic reviews*, 9, 1–124.

Baird, S., C. McIntosh, and B. Özler (2011): "Cash or Condition? Evidence from a Cash Transfer Experiment," *The Quarterly journal of economics*, 126, 1709–1753.

Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, and M. Sulaiman (2018): *Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa*, World Bank.

Banerjee, A., E. Duflo, A. Finkelstein, L. F. Katz, B. A. Olken, and A. Sautmann (2020): "In Praise of Moderation: Suggestions for the Scope and Use of Pre-Analysis Plans for RCTs in Economics," *NBER Working Paper 26993*.

Bank, W. (2015): "Policy Brief: Zambia," *World Bank*.

——— (2018): "World Bank Education Overview: Girls' Education," *World Bank*.

Barrera-Osorio, F., M. Bertrand, L. L. Linden, and F. Perez-Calle (2011): "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia," *American Economic Journal: Applied Economics*, 3, 167–95.

Barro, R. J. and J. W. Lee (2013): "A New Data Set of Educational Attainment in the World, 1950-2010," *Journal of Development Economics*, 104, 184–198.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009): "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly journal of economics*, 124, 1497–1540.

Bellei, C. (2009): "Does Lengthening the School Day Increase Students' Academic Achievement? Results from a Natural Experiment in Chile," *Economics of Education Review*, 28, 629–640.

Benjamini, Y. and Y. Hochberg (1995): "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.

Birdsall, N. et al. (1985): "Public inputs and child schooling in Brazil," *Journal of Development Economics*, 18, 67–86.

Borker, G. (2018): "Safety First: Perceived Risk of Street Harassment and Educational Choices of Women," *Working Paper*.

Burde, D. and L. L. Linden (2013): "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools," *American Economic Journal: Applied Economics*, 5, 27–40.

Cannonier, C. and N. Mocan (2018): "The Impact of Education on Women's Preferences for Gender Equality: Evidence from Sierra Leone," *Journal of Demographic Economics*, 84, 3–40.

Chaudhury, N. and D. Parajuli (2010): "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Programme on Public School Enrolments in Punjab, Pakistan," *Applied Economics*, 42, 3565–3583.

Cohen, J., P. Dupas, et al. (2010): "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment," *Quarterly Journal of Economics*, 125, 1–45.

Cronin, J., G. G. Kingsbury, M. S. McCall, and B. Bowe (2005): "The Impact of the No Child Left Behind Act on Student Achievement and Growth: 2005 Edition." *Northwest Evaluation Association*.

CSO, Z. (1991-2017): "Living Conditions Statistics," *Central Statistical Office of Zambia*.

Das, J., S. Dercon, J. Habyarimana, P. Krishnan, K. Muralidharan, and V. Sundararaman (2013): "School Inputs, Household Substitution, and Test Scores," *American Economic Journal: Applied Economics*, 5, 29–57.

Dehejia, R., C. Pop-Eleches, and C. Samii (2019): "From Local to Global: External Validity in a Fertility Natural Experiment," *Journal of Business & Economic Statistics*, 1–27.

DeTray, D. (1988): "Government Policy, Household Behavior, and the Distribution of Schooling: A Case Study of Malaysia," *Research in Population Economics*, 6, 303–336.

Dhar, D., T. Jain, and S. Jayachandran (2018): "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India," *NBER Working Paper No. 25331*.

DHS (2002): "Zambia DHS EdData Survey," *Central Statistical Office (Zambia), Ministry of Health (Zambia), Macro International, Inc.*, Washington DC, United States: United States Agency for International Development (USAID).

Duflo, E. (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Eeview*, 91, 795–813.

——— (2012): "Women Empowerment and Economic Development," *Journal of Economic Literature*, 50, 1051–79.

Evans, D. K., M. Akmal, and P. Jakiela (2020): "Gender Gaps in Education: The Long View," *Center for Global Development Working Paper 523*.

Fiala, N., A. Garcia-Hernandez, K. Narula, and N. Prakash (2020): "Wheels of Change: Transforming Girl's Lives with Bicycles," *Working Paper*.

Field, E. M., R. Pande, N. Rigol, S. G. Schaner, and C. T. Moore (2019): "On Her Own Account: How Strengthening Women's Financial Control Affects Labor Supply and Gender Norms," *NBER Working Paper No. 26294*.

Filmer, D. and N. Schady (2011): "Does More Cash in Conditional Cash Transfer Programs Always Lead to Larger Impacts on School Attendance?" *Journal of Development Economics*, 96, 150–157.

Fiszbein, A. and N. R. Schady (2009): *Conditional Cash Transfers: Reducing Present and Future Poverty*, The World Bank.

Gechter, M. (2015): "Generalizing the Results from Social Experiments: Theory and Evidence from Mexico and India," *Working Paper*.

Glennerster, R., C. Walsh, and L. Diaz-Martin (2018): "A Practical Guide to Measuring Women's and Girls' Empowerment in Impact Evaluations," *J-PAL*.

Glewwe, P. and K. Muralidharan (2016): "Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications," *Handbook of the Economics of Education, ed. Eric Hanushek, Stephen Machin and Ludger Woessmann*, 653–744.

Haushofer, J. and J. Shapiro (2016): "The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya," *The Quarterly Journal of Economics*, 131, 1973–2042.

Hawke, A. (2015): "Fixing the Broken Promise of Education for All: Findings from the Global Initiative on Out-of-School Children," *UNICEF*.

Heath, R. and A. M. Mobarak (2015): "Manufacturing Growth and the Lives of Bangladeshi Women," *Journal of Development Economics*, 115, 1–15.

Holm, S. (1979): "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, 6, 65–70.

Jayachandran, S. (2015): "The Roots of Gender Inequality in Developing Countries," *Annual Review of Economics*, 7, 63–88.

Kabeer, N. (1999): "Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment," *Development and Change*, 30, 435–464.

Kaffenberger, M., L. Pritchett, and J. Sandefur (2018): "Estimating the Impact of Women's Education on Fertility, Child Mortality, and Empowerment When Schooling Ain't Learning," *Working Paper*.

Kazianga, H., D. Levy, L. L. Linden, and M. Sloan (2013): "The effects of "Girl-Friendly" Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso," *American Economic Journal: Applied Economics*, 5, 41–62.

Khanna, G. (2019): "Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India," *Working Paper*.

Kowalski, A. (2019): "How to Examine External Validity Within an Experiment," *NBER Working Paper 24834*.

La Ferrara, E., A. Chong, and S. Duryea (2012): "Soap Operas and Fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 4, 1–31.

Langton, L. and J. Truman (2014): "Socio-Emotional Impact of Violent Crime," *Washington, DC: Bureau of Justice Statistics*, NCJ 247076.

Laszlo, S. and K. Grantham (2017): "Measurement of Women's Economic Empowerment in GrOW Projects: Inventory and User Guide," *McGill University GrOW Working Paper*, 124–140.

Lavy, V. (1996): "School Supply Constraints and Children's Educational outcomes in Rural Ghana," *Journal of Development Economics*, 51, 291–314.

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

Lillard, L. A. and R. J. Willis (1994): "Intergenerational Educational Mobility: Effects of Family and State in Malaysia," *Journal of Human Resources*, 1126–1166.

Lundberg, S. and R. A. Pollak (1993): "Separate Spheres Bargaining and the Marriage Market," *Journal of Political Economy*, 101, 988–1010.

Macmillan, R., A. Nierobisz, and S. Welsh (2000): "Experiencing the Streets: Harassment and Perceptions of Safety among Women," *Journal of Research in Crime and Delinquency*, 37, 306–322.

Macy, S. (2011): "Wheels of Change: How Women Rode the Bicycle to Freedom (With a Few Flat Tires Along the Way)," *National Geographic Books*.

McConnell, C. and M. Mupuwaliywa (2015): "Keeping Girls in School: Situation Analysis for Zambia," *World Bank*.

Morgan, G. G. (2010): "Handbook of Research on Nonprofit Economics and Management - edited by Bruce A. Seaman and Dennis R. Young," *Public Administration*, 88, 1134–1135.

Muralidharan, K. and N. Prakash (2017): "Cycling to School: Increasing Secondary School Enrollment for Girls in India," *American Economic Journal: Applied Economics*, 9, 321–50.

Mwanza, P. (2015): "The State of Girl-Child's Education in Zambia: The Case of Chongwe District," *Journal of International Cooperation in Education*, 17, 95–110.

Oster, S. M. (1995): *Strategic Management for Non-Profit Organizations: Theory and Cases*, Oxford University Press.

Rawlings, L. B. and G. M. Rubio (2005): "Evaluating the Impact of Conditional Cash Transfer Programs," *The World Bank Research Observer*, 20, 29–55.

Riley, E. (2017): "Increasing Students' Aspirations: The Impact of Queen of Katwe on Students' Educational Attainment," *CSAE Working Paper 2017-13*.

Rosenzweig, M. R. and C. Udry (2020): "External Validity in a Stochastic World: Evidence from Low-Income Countries," *The Review of Economic Studies*, 87, 343–381.

Saito, M. (2011): "Trends in the Magnitude and Direction of Gender Differences in Learning Outcomes," *SACMEQ: UNESCO International Institute for Educational Planning*.

Samarakoon, S. and R. A. Parinduri (2015): "Does Education Empower Women? Evidence from Indonesia," *World Development*, 66, 428–442.

Siddique, Z. (2018): "Violence and Female Labor Supply," *IZA Discussion Paper No. 11874*.

Thaler, R. (1980): "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior & Organization*, 1, 39–60.

UNESCO (2016): "Zambia: Education Policy Review; Paving the way for SDG 4-Education 2030," *UNESCO*.

UNICEF (2014): "Global Initiative on Out-of-School Children," *Republic of Zambia, UNICEF, UNESCO*.

Vivalt, E. (2019): "Specification Searching and Significance Inflation Across Time, Methods and Disciplines," *Oxford Bulletin of Economics and Statistics*, 81, 797–816.

Westfall, P. H., S. S. Young, and S. P. Wright (1993): "On Adjusting P-Values for Multiplicity," *Biometrics*, 49, 941–945.

Main Tables

Table 1: First Stage Outcomes

Dependent variable:	Days Late	Bike Access	Time to school
	(1)	(2)	(3)
		Panel: A	
Pooled Treatment	-1.45***	0.88***	-34.82***
	(0.10)	(0.02)	(2.94)
W-Y p-value	0.00	0.00	0.00
B-H p-value	0.00	0.00	0.00
R-squared	0.20	0.79	0.13
		Panel: B	
Payment Arm	-1.39***	0.89***	-36.44***
	(0.11)	(0.03)	(3.33)
No Payment Arm	-1.53***	0.88***	-33.17***
	(0.10)	(0.02)	(4.04)
Observations	1952	2001	1879
W-Y p-value (Payment Arm)	0.00	0.00	0.00
B-H p-value (Payment Arm)	0.00	0.00	0.00
W-Y p-value (No Payment Arm)	0.00	0.00	0.00
B-H p-value (No Payment Arm)	0.00	0.00	0.00
R-squared	0.20	0.80	0.13
Control group mean	2.19	0.02	103.77
Payment Arm = No Payment Arm (p-value)	0.13	0.92	0.45

NOTES: Standard errors clustered at the school level. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). The dependent variable in (Column 1) is the number of days in the last week the girl arrived late to school, in (2) is a dummy indicating whether the girl has access to a bicycle, and in (3) is the time spent traveling to school (in minutes) each way. W-Y p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Westfall et al. (1993), and B-H p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Holm (1979). *** 1%, ** 5%, * 10%.

Table 2: Impact on Educational Outcomes

Dependent variable:	Days Absent	Dropouts	Grade Transition	Mathematics	English
	(1)	(2)	(3)	(4)	(5)
			Panel: A		
Pooled Treatment	-0.28***	-0.02	-0.03	0.11*	0.03
	(0.08)	(0.02)	(0.02)	(0.06)	(0.05)
W-Y p-value	0.00	0.64	0.73	0.42	0.73
B-H p-value	0.00	0.99	0.61	0.89	1.00
R-squared	0.04			0.34	0.35
			Panel: B		
Payment Arm	-0.27***	-0.03	-0.02	0.13*	0.06
	(0.09)	(0.02)	(0.02)	(0.07)	(0.06)
No Payment Arm	-0.27***	-0.01	-0.03	0.07	-0.02
	(0.10)	(0.02)	(0.02)	(0.06)	(0.07)
Observations	1952	2448	1931	2001	2001
W-Y p-value (Payment Arm)	0.00	0.64	0.73	0.42	0.73
B-H p-value (Payment Arm)	0.00	0.73	0.92	0.42	0.92
W-Y p-value (No Payment Arm)	0.01	0.99	0.61	0.89	1.00
B-H p-value (No Payment Arm)	0.01	1.00	0.85	1.00	1.00
R-squared	0.04			0.34	0.35
Control group mean	1.01	0.06	0.94	0.00	0.00
Payment Arm = No Payment Arm (p-value)	0.98	0.43	0.69	0.44	0.30

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). The dependent variable in Column (1) is the number of days the girl missed school in the last week, in Column (2) is a dummy = 1 if the girl dropped out of school, in (3) is a dummy = 1 if the girl progressed to a higher grade, conditional on not dropping out, in (4) and (5) is the standardized learning assessment score in Mathematics and English self administered test. W-Y p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Westfall et al. (1993), and B-H p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Holm (1979). *** 1%, ** 5%, * 10%.

Table 3: Medium Run Impact on Girls Attendance in 2019

Dependent variable:	Days Absent in 2019	
	Term I	Term II
	(1)	(2)
	Panel: A	
Pooled Treatment	-3.27***	-3.13***
	(0.78)	(0.74)
	Panel: B	
Payment Arm	-2.91***	-2.82***
	(0.99)	(0.94)
No Payment Arm	-3.69***	-3.51***
	(0.75)	(0.77)
Observations	1469	1467
R-squared	0.10	0.09
Control group mean	8.37	8.44
Payment Arm = No Payment Arm (p-value)	0.38	0.45

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and grade fixed effects. The dependent variable in Column (1) and (2) is number of days missed by a girl in a term in 2019. All the columns use administrative data from the school registers collected by the District Educational Board Secretaries and coordinated by the World Bicycle Relief. *** 1%, ** 5%, * 10%.

Table 4: Medium Run Impact on Girls Educational Outcomes in 2019

Dependent variable:	Enrollment	Dropout	Grade Transition
	(1)	(2)	(3)
		Panel: A	
Pooled Treatment	0.07**	-0.07**	0.02
	(0.03)	(0.03)	(0.02)
		Panel: B	
Payment Arm	0.06*	-0.06*	0.04*
	(0.04)	(0.04)	(0.02)
No Payment Arm	0.08**	-0.08**	0.00
	(0.04)	(0.04)	(0.03)
Observations	2467	2467	1808
R-squared	0.05	0.05	0.03
Control group mean	0.63	0.37	0.75
Payment Arm = No Payment Arm (p-value)	0.42	0.42	0.83

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and grade fixed effects. The dependent variable is a dummy variable that indicates whether the girl is still in enrolled in school in the third term of 2019. All the columns use administrative data from the school registers collected by the District Educational Board Secretaries and coordinated by the World Bicycle Relief. *** 1%, ** 5%, * 10%.

Table 5: Impact on Measures of Empowerment

Dependent variable:	Mobility & Safety	Aspirations	Control	Fertility & Marriage
	(1)	(2)	(3)	(4)
			Panel: A	
Pooled Treatment	0.00	0.06	0.16***	0.10
	(0.06)	(0.05)	(0.06)	(0.07)
W-Y p-value	0.94	0.37	0.16	0.37
B-H p-value	1.00	1.00	0.46	1.00
R-squared	0.01	0.02	0.02	0.07
			Panel: B	
Payment Arm	-0.01	0.11**	0.18**	0.18**
	(0.07)	(0.05)	(0.07)	(0.09)
No Payment Arm	0.02	-0.00	0.13*	-0.01
	(0.07)	(0.08)	(0.07)	(0.08)
Observations	1935	1919	2005	1945
W-Y p-value (Payment Arm)	0.94	0.37	0.16	0.37
B-H p-value (Payment Arm)	0.92	0.33	0.16	0.33
W-Y p-value (No Payment Arm)	1.00	1.00	0.46	1.00
B-H p-value (No Payment Arm)	1.00	1.00	0.55	1.00
R-squared	0.01	0.03	0.02	0.07
Control group mean	0.33	1.44	0.50	0.95
Payment Arm = No Payment Arm (p-value)	0.73	0.16	0.55	0.08

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). The dependent variable in Column (1) is the index on mobility, in Column (2) is the index on aspirations, in Column (3) is the index on locus of control, and in Column (4) is the index on fertility and marriage. All indices have been variance-weighted using the methodology of Anderson (2008). Endline indices contain imputed values if less than 10% of the variables in the index had missing values for an observation. The p-values adjusted for multiple hypotheses testing are corrected with the outcome variables in Table XX of the main outcomes in empowerment. W-Y p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Westfall et al. (1993), and B-H p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Holm (1979). *** 1%, ** 5%, * 10%.

Table 6: Impact on Behavioral Outcomes

Dependent variable:	Bargaining	Pro-Sociality	Self-Image
	(1)	(2)	(3)
		Panel: A	
Pooled Treatment	0.19***	0.14**	0.11*
	(0.05)	(0.06)	(0.06)
W-Y p-value	0.05	0.45	0.41
B-H p-value	0.14	0.46	0.85
R-squared	0.05	0.05	0.04
		Panel: B	
Payment Arm	0.18***	0.11*	0.13*
	(0.06)	(0.06)	(0.07)
No Payment Arm	0.18***	0.16*	0.09
	(0.07)	(0.08)	(0.07)
Observations	1988	1874	1889
W-Y p-value (Payment Arm)	0.05	0.45	0.41
B-H p-value (Payment Arm)	0.06	0.42	0.36
W-Y p-value (No Payment Arm)	0.14	0.46	0.85
B-H p-value (No Payment Arm)	0.08	0.55	1.00
R-squared	0.05	0.05	0.04
Control group mean	0.40	0.77	0.68
Payment Arm = No Payment Arm (p-value)	0.98	0.58	0.59

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). The dependent variable in Columns (1) and (2) is an index of bargaining, in Columns (3) and (4) is an index of pro-sociality, and in Columns (5) and (6) is an index of self-image. All indices have been variance-weighted using the methodology of Anderson (2008). Endline indices contain imputed values if less than 10% of the variables in the index had missing values for an observation. The p-values adjusted for multiple hypotheses testing are corrected with the outcome variables in Table XX of the main outcomes in empowerment. W-Y p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Westfall et al. (1993), and B-H p-value gives the p-value adjusted for multiple hypothesis tests using the methodology of Holm (1979). *** 1%, ** 5%, * 10%.

Table 7: Impact on Safety Outcomes

Dependent variable:	Sub-Index of			
	Mobility	Safety	Teased	Missed School
	(1)	(2)	(3)	(4)
			Panel: A	
Pooled Treatment	-0.02	0.10*	-0.08***	-0.06***
	(0.06)	(0.05)	(0.03)	(0.02)
			Panel: B	
Payment Arm	-0.03	0.12*	-0.08***	-0.06***
	(0.08)	(0.06)	(0.03)	(0.02)
No Payment Arm	-0.01	0.09	-0.08**	-0.06**
	(0.07)	(0.06)	(0.04)	(0.02)
Observations	1989	1936	1954	1953
R-squared	0.01	0.02	0.03	0.01
Control group mean	0.25	0.31	0.37	0.18
Payment Arm = No Payment Arm (p-value)	0.81	0.68	0.90	0.80

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). The dependent variable in Columns (1) and (2) is the sub-index of mobility, in Columns (3) and (4) is the sub-index of perceived safety, in Columns (5) and (6) is the probability of the girl being teased or whistled at on the way to school, and in Columns (7) and (8) is the probability that a girl misses school or leaves early for home for safety concerns. All indices have been variance-weighted using the methodology of Anderson (2008). Endline indices contain imputed values if less than 10% of the variables in the index had missing values for an observation. *** 1%, ** 5%, * 10%.

Table 8: Impact on Time Use

Dependent variable:	School Chores	ECA	Home Work	Household Chores	Income Generation	Friends
	(1)	(2)	(3)	(4)	(5)	(6)
				Panel: A		
Pooled Treatment	0.77	1.03	0.93	1.02	0.75**	1.04
	(0.14)	(0.13)	(0.11)	(0.12)	(0.10)	(0.11)
				Panel: B		
Payment Arm	0.91	1.15	1.02	1.01	0.78	1.06
	(0.20)	(0.17)	(0.15)	(0.16)	(0.13)	(0.15)
No Payment Arm	0.62**	0.89	0.81	1.02	0.73**	0.99
	(0.13)	(0.15)	(0.11)	(0.14)	(0.11)	(0.12)
Observations	1938	1925	1931	1997	2000	2005
Payment Arm = No Payment Arm (p-value)	0.15	0.17	0.14	0.95	0.72	0.66

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). All columns report odds ratio from an ordered logit model. The dependent variable is the time spent by the girl doing various activities, and takes the following values: 0 - no time spent, 1 - less than 30 minutes, 2 - between 30 and 60 minutes, 3 - between 60 and 90 minutes, 4 - between 90 and 120 minutes, 5 - more than 120 minutes. The activities are: School chores like cleaning in column (1), Extra curricular activities (ECA) in column (2), Studying and homework in column (3), Household chores in column (4), Engaging in income generating activities in column (5), and Spending time with friends in Column (6). *** 1%, ** 5%, * 10%.

Table 9: Heterogeneous Impact on Educational Outcomes by Baseline Time Taken to School

Dependent variable:	Bike Access	Days Absent	Time to school	Days Late	Mathematics	English
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.87***	-0.21**	-28.20***	-1.27***	0.12	0.01
	(0.03)	(0.11)	(4.21)	(0.14)	(0.08)	(0.08)
Treatment X Middle Tercile	0.02	-0.27*	-9.13*	-0.29*	0.01	0.09
	(0.03)	(0.14)	(5.38)	(0.16)	(0.08)	(0.08)
Treatment X Top Tercile	0.00	0.13	-15.37**	-0.35*	-0.08	-0.09
	(0.03)	(0.19)	(6.08)	(0.18)	(0.10)	(0.10)
Middle tercile	-0.00	0.14	12.17**	0.14	-0.07	-0.13**
	(0.01)	(0.11)	(5.27)	(0.15)	(0.05)	(0.06)
Top tercile	-0.00	-0.04	6.16	0.07	-0.00	-0.05
	(0.02)	(0.13)	(7.57)	(0.15)	(0.07)	(0.08)
Observations	2001	1952	1879	1952	2001	2001
R-squared	0.79	0.04	0.14	0.20	0.34	0.35
Control group mean (Bottom Tercile)	0.02	0.96	94.00	2.12	0.06	0.07

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). The dependent variable in Column (1) is a dummy for whether the girl has access to a bicycle, in Column (2) is the number of days the girl missed school in the last week, in Column (3) is the time spent travelling to school (in minutes) each way, in Column (4) is the number of days in the last week the girl arrived late to school, in Column (5) is the standardised learning assessment score in Mathematics, and in Column (6) is the standardised learning assessment score in English. *** 1%, ** 5%, * 10%.

Table 10: Heterogeneous Impact on Educational Outcomes by Baseline Time Taken to School

Dependent variable:	Index of						
	Mobility	Aspiration	Locus of Control	Fertility & Marriage	Bargaining	Pro-Sociality	Self-Image
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	-0.01	0.09	0.19**	0.11	0.29***	0.08	0.14
	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)	(0.06)	(0.09)
Treatment X Middle Tercile	0.04	-0.05	0.05	-0.07	-0.08	0.15	0.03
	(0.12)	(0.10)	(0.11)	(0.11)	(0.10)	(0.09)	(0.12)
Treatment X Top Tercile	0.01	-0.03	-0.20*	0.08	-0.29**	0.03	-0.16
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.11)	(0.12)
Middle Tercile	-0.06	0.11	-0.02	-0.08	0.08	-0.06	-0.04
	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.09)
Top Tercile	-0.01	0.18**	0.15*	-0.03	0.22**	0.02	0.14*
	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.08)
Observations	1935	1919	2005	1945	1988	1874	1889
R-squared	0.01	0.03	0.03	0.07	0.06	0.06	0.04
Control group mean (Bottom Tercile)	0.36	1.36	0.46	1.02	0.35	0.80	0.66

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). The dependent variable in Column (1) is the index on mobility and safety, in Column (2) is the index on aspirations, in Column (3) is the index on locus of control, in Column (4) is the index on fertility and marriage, in Column (5) is the index of bargaining (not including the variables on rebellion), in Column (6) is the index of pro sociality, and in Column (7) is the index of self image. *** 1%, ** 5%, * 10%.

Appendix Tables

Table A.1: Balance Table: School Characteristics

	Treatment Groups			p value for test of:		
	Control (N = 55)	No Payment Arm (T2) (N = 20)	Payment Arm (T1) (N = 25)	$1 = 2$	$1 = 3$	$1 = (2 \cup 3)$
	(1)	(2)	(3)	(4)	(5)	(6)
Enrollment 2017	692.75	643.40	686.92	0.31	0.90	0.51
	(187.74)	(186.96)	(192.46)			
Enrollment girls 2017	344.85	318.70	338.04	0.28	0.77	0.41
	(93.89)	(91.98)	(95.58)			
Enrollment boys 2017	347.89	324.70	348.88	0.36	0.97	0.62
	(96.21)	(98.60)	(99.31)			
# teachers	13.47	13.45	12.72	0.99	0.54	0.72
	(6.46)	(6.44)	(4.28)			
% teachers living school	78.48	71.98	68.97	0.39	0.16	0.15
	(27.26)	(29.82)	(28.53)			
Km closest town	51.02	38.92	45.28	0.18	0.50	0.30
	(57.37)	(19.19)	(17.77)			
Km closest tarmac road	25.06	24.63	25.90	0.93	0.84	0.94
	(19.39)	(18.78)	(16.93)			
Km closest secondary school	19.31	19.82	22.10	0.92	0.58	0.64
	(15.76)	(18.78)	(22.02)			
Feeding program	0.04	0.00	0.04	0.16	0.94	0.68
	(0.19)	(0.00)	(0.20)			
Sanitation program	0.60	0.60	0.60	1.00	1.00	1.00
	(0.49)	(0.50)	(0.50)			
Other program	0.35	0.45	0.24	0.42	0.33	0.90
	(0.48)	(0.51)	(0.44)			
# toilets girls	5.78	4.70	4.88	0.09	0.14	0.05
	(2.68)	(2.34)	(2.40)			
# toilets boys	5.24	3.85	4.84	0.01	0.53	0.10
	(2.83)	(1.81)	(2.46)			
# classrooms 2017	8.44	7.95	8.08	0.58	0.59	0.51
	(3.43)	(3.30)	(2.34)			
Access to Library	0.31	0.20	0.16	0.33	0.13	0.13
	(0.47)	(0.41)	(0.37)			
Access to computers	0.95	0.95	0.92	0.94	0.69	0.80
	(0.23)	(0.22)	(0.28)			

NOTES: Balance test of the school characteristics by treatment groups. *** 1%, ** 5%, * 10%.

Table A.2: Balance Table: Girls Characteristics

	Treatment Groups			p-value for test of:		
	Control (N = 1357)	No Payment Arm (T2) (N = 500)	Payment Arm (T1) (N = 614)	1 = 2	1 = 3	1 = (2 ∪ 3)
	(1)	(2)	(3)	(4)	(5)	(6)
Age	12.88	12.96	12.85	0.27	0.62	0.79
	(1.43)	(1.44)	(1.36)			
Grade in school	6.02	6.09	6.07	0.07	0.13	0.04
	(0.82)	(0.82)	(0.81)			
Ever repeated a grade	0.36	0.37	0.36	0.59	0.81	0.66
	(0.48)	(0.48)	(0.48)			
Both parents alive	0.80	0.80	0.83	0.99	0.21	0.47
	(0.40)	(0.40)	(0.38)			
Household size	6.46	6.21	6.39	0.08	0.67	0.23
	(2.88)	(2.54)	(3.23)			
# of biological brothers	1.70	1.79	1.56	0.31	0.06	0.57
	(1.53)	(1.74)	(1.49)			
# of biological sisters	1.36	1.25	1.35	0.12	0.88	0.35
	(1.40)	(1.30)	(1.37)			
Currently engaged/married	0.85	0.84	0.84	0.85	0.76	0.78
	(0.36)	(0.36)	(0.37)			
Ever been pregnant	0.06	0.06	0.05	0.77	0.24	0.36
	(0.24)	(0.23)	(0.22)			
# of meals with Meat	1.62	1.72	1.78	0.24	0.04	0.04
	(1.52)	(1.63)	(1.58)			
# of days with more than 1 meal	5.77	5.78	5.74	0.96	0.74	0.83
	(2.36)	(2.39)	(2.41)			
# of days with not enough food	0.97	0.81	0.92	0.06	0.54	0.14
	(1.67)	(1.46)	(1.68)			
Socio-economic index (PCA)	-0.08	0.04	0.14	0.12	0.02	0.01
	(1.51)	(1.50)	(2.20)			
Locus of control index (PCA)	0.01	-0.14	0.08	0.13	0.45	0.69
	(1.90)	(1.89)	(1.87)			
Self esteem index (PCA)	-0.05	0.05	0.08	0.22	0.07	0.06
	(1.51)	(1.55)	(1.55)			
Gender attitudes index (PCA)	-0.03	0.00	0.07	0.64	0.14	0.23
	(1.36)	(1.34)	(1.34)			

NOTES: Balance test of the school characteristics by treatment groups. *** 1%, ** 5%, * 10%.

Table A.3: Tracking and Attrition

	Total	Control	Payment	No Payment
Tracked and surveyed	2,028	1,071	533	424
In school	1,789	920	481	388
Second stage tracking	239	151	52	36
Attrition Rate (not weighted)	17.9%	21.1%	13.2%	15.2%
Effective Attrition Rate (weighted)	8.7%	10.1%	7.8%	5.6%

NOTES: The sample analyzed in this table are the 2,471 potential endline respondents.

Table A.4: Determinants of Sample Attrition

	Attrited Endline
Payment Arm	-0.013
	(0.026)
No Payment Arm	-0.053*
	(0.022)
Payment Arm X Socio-Economic Index (PCA)	-0.007
	(0.010)
No Payment Arm X Socio-Economic Index (PCA)	-0.007
	(0.011)
Payment Arm X Locus of Control Index (PCA)	-0.008
	(0.012)
No Payment Arm X Locus of control Index (PCA)	-0.000
	(0.009)
Payment Arm X Self-Esteem Index (PCA)	-0.006
	(0.012)
No Payment Arm X Self-Esteem Index (PCA)	0.002
	(0.015)
Payment Arm X Gender Attitudes Index (PCA)	0.015
	(0.015)
No Payment Arm X Gender Attitudes Index (PCA)	0.018
	(0.014)
Socio-Economic Index (PCA)	0.008
	(0.007)
Locus of Control Index (PCA)	0.001
	(0.008)
Self-Esteem Index (PCA)	0.001
	(0.006)
Gender Attitudes Index (PCA)	0.001
	(0.009)
Control group mean	0.101
Payment Arm mean	0.089
No Payment Arm mean	0.049
Payment Arm = Control (p-value)	0.632
No Payment Arm = Control (p-value)	0.016
Payment Arm = No Payment Arm (p-value)	0.054
Observations	2,467

NOTES: Standard errors in parentheses and clustered at the school level. This includes the survey weights used because of the two-stage sampling procedure. *** 1%, ** 5%, * 10%.

Table A.5: Lee Bounds: Educational Outcomes

Dependent variable:	Days Absent			Days Late			Time to School			Mathematics			English		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Pooled Treatment	-0.28***	-0.43***	-0.28***	-1.45***	-1.55***	-1.45***	-34.82***	-41.99***	-32.79***	0.11*	0.08	0.15***	0.03	-0.03	0.06
	(0.08)	(0.08)	(0.08)	(0.10)	(0.10)	(0.10)	(2.94)	(2.36)	(2.92)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
Observations	1952	1923	1952	1952	1927	1952	1879	1845	1841	2001	1982	1973	2001	1963	1986
R-squared	0.04	0.06	0.04	0.20	0.24	0.20	0.13	0.22	0.12	0.34	0.33	0.34	0.35	0.30	0.35
Control group mean	1.01	1.01	1.01	2.19	2.19	2.19	103.77	103.70	103.77	0.00	0.00	0.00	0.00	0.00	0.00
Lee Bounding	No	Upper	Lower	No	Upper	Lower	No	Upper	Lower	No	Upper	Lower	No	Upper	Lower

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). We follow Lee (2009) and the following two assumptions to do the Lee Bounding: (i) Random assignment of treatment, and (ii) Monotonicity assumption about selection mechanism (treatment affects attrition only in one direction, and girls would have attrited if in control but girl does not attrit because of receiving the bicycle). We calculated the proportion of sample to trim with the following formula:

$$pTrim = (\frac{0.93 - 0.89}{0.93}) * 100 = 4.3$$

The tracking rate of the control group = 89.9, and the tracking rate of the pooled treatment = 93. Sample was trimmed such that the share of observed girls is equal for both groups (we found more girls in the treatment groups than control group, therefore, we trim the treatment group). *** 1%, ** 5%, * 10%.

Table A.6: Lee Bounds: Empowerment Outcomes

Dependent variable:	Mobility & Safety			Aspirations			Control			Fertility		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Pooled Treatment	0.00	-0.08	0.10*	0.06	0.06	0.20***	0.16***	0.16***	0.26***	0.10	0.02	0.24***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.05)
Observations	1935	1896	1897	1919	1919	1881	2005	2005	1965	1945	1905	1906
R-squared	0.01	0.01	0.01	0.02	0.02	0.03	0.02	0.02	0.04	0.07	0.06	0.08
Control group mean	0.33	0.33	0.33	1.44	1.44	1.44	0.50	0.50	0.50	0.95	0.95	0.95
Lee Bounding	No	Upper	Lower	No	Upper	Lower	No	Upper	Lower	No	Upper	Lower

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). All columns report coefficients from a linear model. The dependent variable in (1-3) is the index on mobility, in (4-6) is the index on aspirations, in (7-9) is the index on locus of control, and in (10-12) is the index on fertility and marriage. All indices have been variance-weighted using the methodology of Anderson (2008). Endline indices contain imputed values if less than 10% of the variables in the index had missing values for an observation. We follow Lee (2009) and the following two assumptions to do the Lee Bounding: (i) Random assignment of treatment, and (ii) Monotonicity assumption about selection mechanism (treatment affects attrition only in one direction, and girls would have attrited if in control but girl does not attrit because of receiving the bicycle). We calculated the proportion of sample to trim with the following formula:

$$pTrim = (\frac{0.93 - 0.89}{0.93}) * 100 = 4.3$$

The tracking rate of the control group = 89.9, and the tracking rate of the pooled treatment = 93. Sample was trimmed such that the share of observed girls is equal for both groups (we found more girls in the treatment groups than control group, therefore, we trim the treatment group). *** 1%, ** 5%, * 10%.

Table A.7: Lee Bounds: Behavioral Outcomes

Dependent variable:	Bargaining			Pro-Sociality			Self-Image		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pooled Treatment	0.19***	0.19***	0.31***	0.14**	0.08	0.23***	0.11*	0.11*	0.20***
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Observations	1988	1988	1948	1874	1841	1843	1889	1889	1851
R-squared	0.05	0.05	0.07	0.05	0.05	0.06	0.04	0.04	0.04
Control group mean	0.40	0.40	0.40	0.77	0.77	0.77	0.68	0.68	0.68
Lee Bounding	No	Upper	Lower	No	Upper	Lower	No	Upper	Lower

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). All columns report coefficients from a linear model. The dependent variable in (1-3) is the index of bargaining (not including the variables on rebellion), in (4-6) is an index of pro-sociality, and in (7-9) an index of self-image. All indices have been variance-weighted using the methodology of Anderson (2008). Endline indices contain imputed values if less than 10% of the variables in the index had missing values for an observation. We follow Lee (2009) and the following two assumptions to do the Lee Bounding: (i) Random assignment of treatment, and (ii) Monotonicity assumption about selection mechanism (treatment affects attrition only in one direction, and girls would have attrited if in control but girl does not attrit because of receiving the bicycle). We calculated the proportion of sample to trim with the following formula:

$$pTrim = (\frac{0.93 - 0.89}{0.93}) * 100 = 4.3$$

. The tracking rate of the control group = 89.9, and the tracking rate of the pooled treatment = 93. Sample was trimmed such that the share of observed girls is equal for both groups (we found more girls in the treatment groups than control group, therefore, we trim the treatment group). *** 1%, ** 5%, * 10%.

Table A.8: Minimum Detectable Effects

Outcome Variables	MDE	N	Mean	N	Mean	N	Mean	ICC
	Pooled	Control	Control	No Payment Arm	No Payment Arm	Payment Arm	Payment Arm	
# Days absent	0.162	1029	1.013	410	0.715	517	0.683	0.072
# Days Late	0.162	1029	2.188	410	0.566	517	0.704	0.072
Time to school	0.154	997	103.774	386	68.422	500	63.916	0.065
Drop-outs	0.124	1356	0.065	483	0.052	613	0.039	0.042
Grade Transition	0.203	1015	0.937	403	0.893	515	0.918	0.114
Access	0.169	1061	3.022	414	3.94	530	3.928	0.078
Index of Mobility	0.176	1023	0.327	404	0.376	514	0.347	0.085
Index of Aspiration	0.137	1014	1.442	403	1.431	506	1.591	0.051
Index of Locus of Control	0.161	1064	0.502	419	0.664	526	0.704	0.072
Index of Fertility	0.201	1039	0.954	420	0.981	518	1.173	0.113
Index of Bargaining	0.183	1050	0.399	419	0.619	523	0.611	0.093
Index of Pro-Sociality	0.16	987	0.772	395	0.945	496	0.909	0.07
Index of Self-image	0.158	997	0.68	397	0.79	499	0.832	0.069

NOTES: This table reports MDEs at $(1-\beta) = 0.8$ and $\alpha = 0.05$. Define μ_i as the mean outcome in arm $i \in \{1, 2, c\}$, then the MDE relies on a t-test with $H_0 : \frac{\mu_1+\mu_2}{2} - \mu_{control} = 0$, i.e. the program has an overall effect.

Table A.9: Descriptive Statistics

	Mean	Std.Dev.	Observations
Demographics			
Age	12.89	1.42	2461
Grade in school	6.05	0.82	2469
Ever repeated a grade	0.36	0.48	2469
Both parents alive	0.81	0.39	2467
Household size	6.39	2.91	2468
# of biological brothers	1.68	1.57	2469
# of biological sisters	1.34	1.37	2469
Currently engaged/married	0.14	0.35	2431
Ever been pregnant	0.06	0.23	2434
Mobility			
Mostly walks to school	0.98	0.13	2467
Time spent traveling to school (mins/each way)	109.16	50.52	2291
Mostly travels to school alone	0.27	0.45	2464
# of people that travel to school together	4.38	3.68	2166
Ever teased on way to school (last year)	0.35	0.48	2469
Would walk to school alone if felt safe	0.79	0.41	2464
Would walk to other places alone if felt safe	0.44	0.50	2459
Attendance			
# of days absent from school (last week)	0.88	1.29	2459
# of days arrived late to school (last week)	2.61	1.69	2412
Learning Assessment			
Learning assessment score (Overall)	0.36	0.16	2468
Learning assessment score (English)	0.30	0.17	2468
Learning assessment score (Maths)	0.44	0.20	2468

NOTES: Descriptive statistics of the girls in the estimation sample.

Table A.10: Impact on Index of Focus (d2 Test)

Dependent variable:	Index of Focus	Speed	Accuracy
	(1)	(2)	(3)
		Panel: A	
Pooled Treatment	0.07	6.11	9.70
	(0.08)	(9.04)	(9.08)
		Panel: B	
Payment Arm	0.15	14.55	18.45
	(0.11)	(11.95)	(12.37)
No Payment Arm	-0.05	-5.73	-2.49
	(0.09)	(9.47)	(9.19)
Observations	1932	1932	1932
R-squared	0.00	0.00	0.00
Control group mean	0.55	459.06	393.45
Payment Arm = No Payment Arm (p-value)	0.12	0.12	0.12

NOTES: Standard errors in parentheses are clustered by school. All regressions include controls for demographics and baseline value of the dependent variable (wherever available). All columns report coefficients from a linear model. The dependent variable in (1) is the index of focus (variance-weighted index of speed and accuracy), in (2) is a measure of speed, which is the total number of observations processed in the d2 test, in (3) is a measure of accuracy, which is the correct number of observations processed in the d2 test. All indices have been variance-weighted using the methodology of Anderson (2008). *** 1%, ** 5%, * 10%.

Appendix Figures

Figure A.1: Map of the Distribution of Schools in the Study Sample

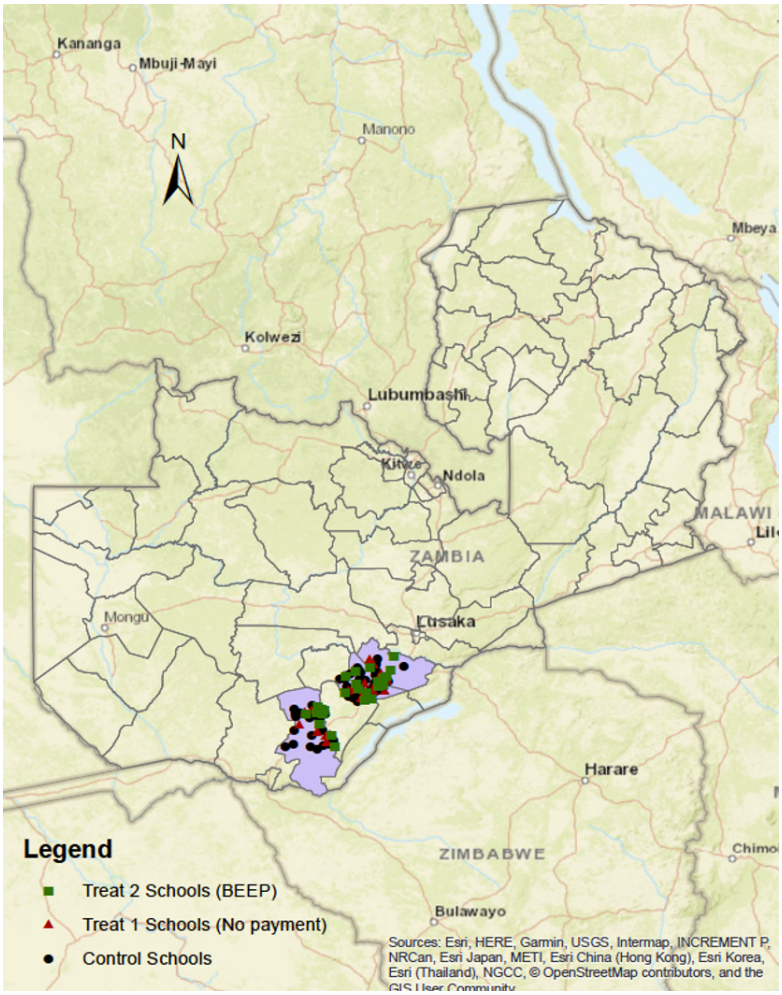


Figure A.2: Timeline

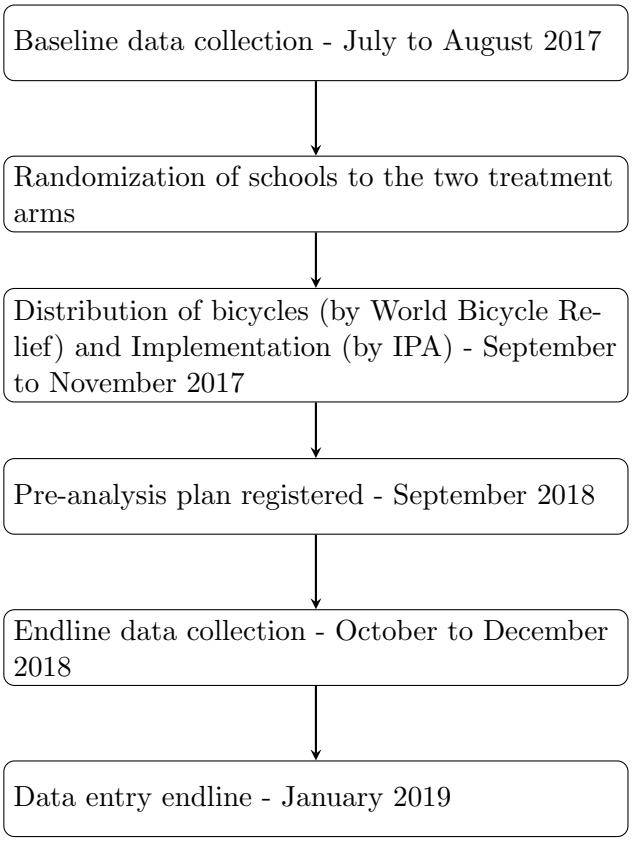
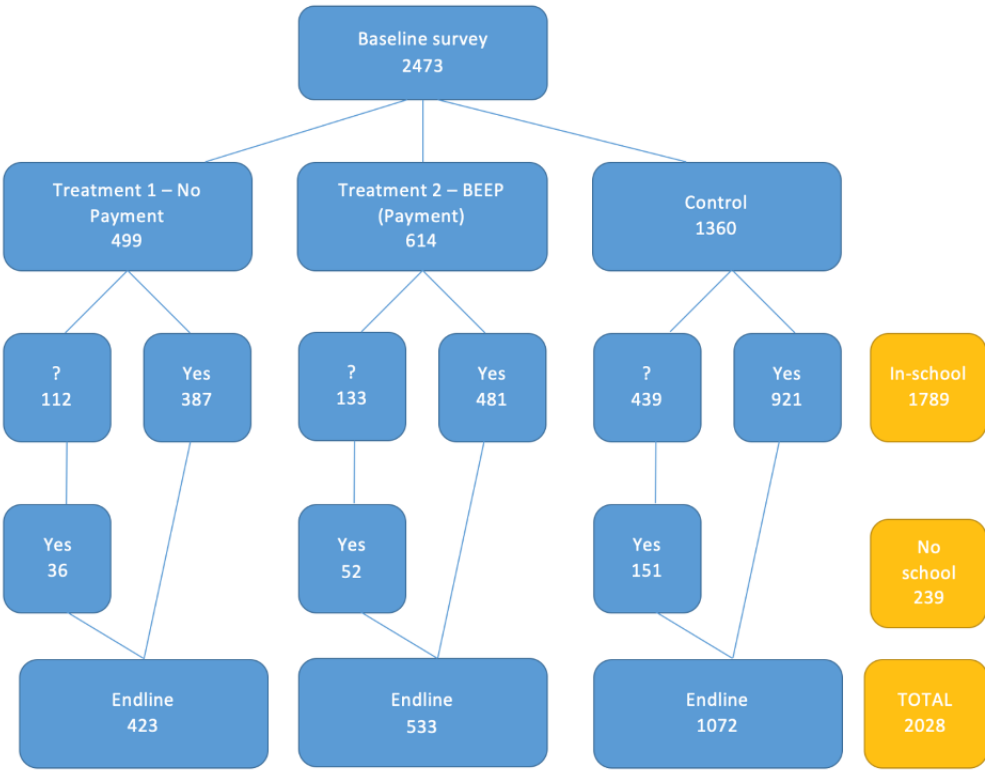


Figure A.3: Tracking and Attrition



A Construction of Outcome Variables

Outcomes	**Variables**
Primary:	
Mobility and Safety	- In the past week, how many times did you go outside the house alone to.. - Are you allowed to go alone when... - I feel safe when...
School Attendance	- Self-reported - Admin data from registers
Grade Transition	- Dropout - Grade Progression
Aspiration	- When you finish at school (either end of primary, secondary), what would you like to do? - How confident are you that you will be able to achieve this? - If for some reason you cannot (insert answer to previous question), what would you do? - In two years of time, how confident are you that you will be enrolled in school? - Do you think you will be working in a job or doing something that makes money in 10 years from now? - I am going to show you some drawings, could you tell me which ones you think a girl like you can become, if any? - And from these same drawings, which one would you like to become when you grow up, if any? - What does your role model do? (Occupation) - Do you want to do what he/she (Role model) does?

Outcomes	**Variables**
Locus of Control	- Let's say that one day when you are going to school you cannot find a path because of heavy rain or because a tree fell. This is a difficult situation because it is the only way to school and you are already late to school. In situations like this one or other ones similar to this one, you can usually find your way out? - There are many things that can happen to you in life. Some of them will be good and some will be not so good. For example: falling over and hurting my knee; forgetting to prepare for an exam or not doing well on an exam; your best friend is upset with you and not talking to you; you were not selected for a school team/club. Do you feel you can control what happens to you in life? - In general, would you say you are satisfied with your life? - I feel my life will improve in the future.
Fertility and Marriage	- The number of children to have in your whole life, how many would you like to have? - Of those children, how many girls and how many boys would you like to have? - Have you ever been pregnant? - Do you have any kids?
Secondary:	
Bicycle ownership and usage	- Do you have access to a bicycle that you can use? - Does this bicycle belong to you? - In a normal week, how many days do you use a bicycle to go to school? - Do you use a bicycle during the weekends? How often? - Do you have to ask permission to use the bicycle? - How much control do you think you have over the bicycle?

Outcomes	**Variables**
Time spent traveling to school	- Working for the school? (doing school chores like cleaning the classes, etc.) - Attending extra-curricular activities? (like sport, production unit, club, drama, board games, etc.) - Studying and doing homework outside of school? - Helping your family at home or doing other work for them? - Working to earn money by yourself? - Being with friends (chatting, playing, games, visiting them at home)?
Performance - Overall score and fraction in the lowest quartile	- Grade 7 end of year exam - English test - Mathematics test - D2 test of Focus
Bargaining	- Do you ever have small money of your own (K2 or K5) to use as you would like? This could be money you have earned or that you get from a family member. - Can you decide on what to spend it on your own? - Each year there are new fashions (e.g. hair pins) that come out. If you wanted to buy something new and had the money to do so, do you think your parents would allow you? - Do you own a pair of leggings? - Do you wear them on their own (if yes)? - If you don't like what is prepared for dinner, would you tell your mother/guardian you don't like the food or ask them if there is something else to eat? - When we talked about the activities you perform at home, like (insert activity here). Have you ever skipped doing household chores?

Outcomes	Variables
	- How often do you say something to your parents if you disagree with what they are saying? - Do you feel you can to talk to your parents about what you want to be when you grow up? - Do you think you can talk to your parents if you have problems with friends or at school? - Do you feel you can talk to your parents about when you wish to get married?
Self-image	- How would you rank yourself academically in your class? - Compared to your friends, how likely are you to succeed in life?
Identity	- Now let's play again with some drawings. Here you can see six drawings of roles girls usually take in society. Can you put them in order, starting from the one you that describes you better to the one that describes you the least? - How much do you think you can affect what other people think of your family?
Pro-sociality	- If you notice that one of your friends has a problem, would you help/participate/collaborate? - Could you tell us the name of your MP? - What is the name of the president of Zambia? - Are you a member of any club? - Think about the most active person in the club and the least active one. The most active would be a 10 and the least active would be a 0. How active are you in this club? - When you don't understand something in class, do you ask the teacher in front of everyone? (Not for out-of-school girls) - Do friends seek your opinion about important matters?

B Sampling Procedure

B.1 School Sampling

We decided to focus only on government schools, which are public schools and also the most common kind in Zambia. In addition, to be able to follow our sample over several years and observe a longer-term impact of the bicycles, all the schools selected are basic schools: Starting at Grade 1 or earlier and going beyond Grade 7 (end of primary) up to Grade 9 (last grade before secondary education).

All basic government schools of Monze and Mazabuka (our initial catchment districts) were asked to identify their pupils walking at least 3 kilometers to school and to generate a list with their names, gender and grade. The research team had to find 100 schools with at least 25 eligible girls enrolled in grade 5, 6 and 7. Many of the schools which prepared the lists didn't have enough of such pupils. Hence, the research team had to extend the catchment area to a third district to find additional candidates for the sample. Kalomo, a third district of Southern Province, was chosen to have a good number of basic government schools, and not much prior work had been done there by World Bicycle Relief. In addition, some schools were also automatically excluded from the sample: (i) urban schools, where the bicycles wouldn't be required by children to travel to school (existence of alternative public transportation), and (ii) a few very remote schools, which created logistical challenges in planning fieldwork. Limited by these constraints, the research team had a limited sample, from which the 100 schools were selected.

B.2 Girls' Sampling

Prior to randomization, the research team had to identify a sample of 25 girls in each school to participate in the data collection activities to satisfy the power calculations. All these girls were required to be enrolled in grades 5, 6 or 7 (grades during which the girls are considered particularly vulnerable and likely to drop out of school).

Among the 100 sample schools, some of them had only 25 eligible girls (girls in grade 5, 6 or 7 and walking 3km or more to come to school), while other schools had 40, 50 or more of such girls. To build a representative sample, we generated two lists for each school. The first list, called list A, would always contain 25 names, balanced across our 3 sample grades (with a small priority given to the grade 7-in which the girls are more likely to drop out of school soon - following as much as possible

the pattern 8-8-9 for the grades 5-6-7). To generate the list, we used Stata and the only variable considered was the grade, like described above.

Then a second list was generated for each school, list B. This second list was containing additional girls, randomly selected among the remaining eligible girls (same methodology). The second list had between 0 and 25 girls, depending on how many girls in total were eligible in the school (in the grades 5, 6 and 7). For example, a school with 33 eligible girls designated in grades 5, 6 and 7 might have a first list of 25 girls to be surveyed (list A), and a second list of 8 girls (list B). If the school had 70 eligible girls in grades 5, 6 and 7, then 25 of them would appear on the list A, 25 others would appear on list B and 20 of them would not appear on any list, the selection being entirely random.

The purpose of these two lists was so that if everything ran smoothly, our field team of surveyors would be able to find the 25 girls of the list A in the school and survey them on the day of their visit. If some of the girls were not able to be surveyed (absent, no consent, transferred, fake name, etc.) then the field team would be able to replace them with girls from the list B. The result would be that the field team would be able to survey a sample of 25 girls in most cases. The names on the list A and B being arranged in a random order, we introduce a limited bias when replacing the names. The only bias introduced was that the field team surveys only those girls who were present. Fortunately, however, those present at the school were not representative of those present on a normal day, because prior to the visit of the team, the girls would have been encouraged by the school to attend on this special day.

C Field Protocols

Several tasks were assigned to the different members of the field team while visiting the schools. The protocol is detailed below:

(i) A few days before the team visits the school (surveyors and supervisors), the school is visited for the first time by a logistics supervisor. The logistics supervisor introduces Innovations for Poverty Action (IPA) and explains to the school staff the involvement of the school in the study. It's important to note that no formal communication from IPA was every made to the schools prior to this visit. To illustrate the legitimacy of the procedure, the logistics supervisor carries two letters, one from the Ministry of General Education (MOGE), one from the District Education Board Secretary (DEBS) both showing the support of these institutions of the study. In a

context where IPA or the Research team is absolutely unknown, much of the involvement from the schools is obtained this way. The logistics supervisor explains that IPA is independent from WBR, although working in collaboration with them, and that participating, or not participating, in the study has no implication on the program that WBR is implementing. Once this is clear, the logistics supervisor provides the lists of girls to be surveyed and asks the school to collect the written informed consent of the parents to let their daughters participate in the study. The logistics supervisor does not speak to every parent whose daughter is asked to participate in the study. Instead, he leaves in the school a pile of consent forms (translated into the local language, Tonga) to be distributed by the school to the parents whose daughter appears on the list, which is also left at the school. A copy of the consent form is to be kept by the parents, another copy is to be signed and returned to the school for the research team. Between this day and the date of the visit of the team in the school, the Logistic supervisors is asked to follow-up by phone or directly on site to check if the consent forms will be ready.

(ii) A few days later (usually between 2 and 5 days), the field team finally arrives in the school. The supervisor meets with the Head-Teacher (or the acting Head) and collects the consent forms which have been signed. Then he/she checks and gathers the first 25 girls from the list, who have the consent of their parents and who are present at school.

(iii) Once the 25 girls are gathered, they are divided in groups of 4 to 5 pupils. Each group goes with one of the surveyors, who is in charge of explaining to them their role in the study. The girls who consent to participate sign an assent form.

(iv) The supervisor visits each of his/her surveyors to give him/her the IDs of the girls who are in his/her group. It's crucial that each surveyor gets the correct IDs because these will enable him/her to connect the face to face interview to the paper based data.

(v) After this, the group activities begin, which are all paper based: the attention test (10 minutes), the learning assessment (25 minutes) and the semi self-administered survey. Each surveyor is supposed to explain and supervise these activities with its own group of 4 or 5 girls. Once the group activities are finished, the girls are released but asked not to go too far. A snack (biscuits and milk drink) is distributed to ease their wait.

(vi) Each surveyor then starts interviewing one of the girls in his/her group, with the other girls waiting some time to be interviewed (they might even go back to class if the surveyor is sure to find them easily again). The face-to-face interview (tablet based) usually lasts around 40

minutes. Once one is finished, the surveyor releases the girl and starts interviewing another girl, until he has interviewed all the girls of his/her group. If the surveyor finishes his/her interviews way before one colleague, he/she can help him/her to finish their interviews.

(vii) Meanwhile, the supervisor conducts the school survey with the head-teacher or the acting head-teacher. Another teacher is welcome to participate if he/she can complete the knowledge of the head. This survey takes between 30 minutes to 1 hour depending on how organized the school is.

(viii) After finishing the school survey, the supervisor asks the school management to prepare the attendance registers to be photographed by the team (those for the grades 5, 6 and 7 for the current and the past years), and all the attendance registers currently available in the school more generally. This is because the surveyors will have to collect in those the attendance of the siblings of the respondents (who are not necessarily enrolled in the same grades).

(ix) When the surveyors finish their interviews, they come to meet with their supervisor and add to their forms the attendance information they collect in the registers (for their respondent and for their siblings).

(x) It's only when all this work is finished, that the team can head back to town and meet with the RA to deliver the data collected and the outputs of the day.

(xi) A few days later, the school might be visited a last time, by one of the back-checkers. Only half of the schools will be back-checked. The back-checker, with no prior notice, will interview again of the girls again, with a short sub-survey (10-15 minutes). No additional consent or assent form needs to be signed.

D Timeline

The baseline data collection happened during the second term of the school year in 2017. It took place between the 5 July and the 10 August 2017, and the team (supervisors and surveyors) worked 21.5 days in the schools (20 days initially planned), over a period of 5 weeks (4 weeks initially planned).

The baseline survey was first launched in Monze and 9 days were necessary to visit all the schools of the district (44 schools). Then, the research team moved to Mazabuka and visited all the schools of

the district (20 schools) in 4 days. Finally, the research team spent 8.5 days in the schools in Kalomo (36 schools).

The training for the supervisors, surveyors, and back-checkers lasted 5 days, including one day of field training. The supervisors had one extra day of training. The training happened in Monze between the 26 June and 1 July 2017.

The logistics supervisors were trained in Lusaka before everyone (1-day training), and they started to visit the schools on 26 June 2017, earlier than the rest of the team, to start planning the visits. The back-checkers finished their work on the same day as most of the rest of the teams.

After one school year using the bicycles, the endline survey was implemented, during the third term of the 2018 school year (September to November 2018).

E Steps for Index Construction

We create variance-weighted indices following the methodology proposed by Anderson (2008) for empowerment outcomes (also see Haushofer and Shapiro (2016); Dhar et al. (2018) for a recent application).

Anderson (2008) summarizes the index creating process as the following. At the most basic level, an index created using this method is a weighted mean of several standardized variables. More weight is assigned to measures that are orthogonal (less similar or less correlated) to other measures. The weights are calculated to maximize the amount of information captured in the index. The index is computed using the following steps.

(i) For all variables, switch signs where necessary so that the positive direction always indicates a "better" outcome.

(ii) Create standardized variables ($\tilde{y}$) by demeaning and then by dividing by standard deviation.

(iii) Compute covariance matrix $\widehat{\sum}$, which consist of elements:

$$\sum_{mn}^{\wedge} = \sum_{i=1}^{Nmn} \frac{(y_{im} - \bar{y}_m)}{\sigma_m^y} * \frac{(y_{in} - \bar{y}_n)}{\sigma_n^y}$$

where, Nmn is the number of observations (total persons with non-missing data for variables m and n).

(iv) Next, we invert the covariance matrix, and define weight w_k for each variable k by summing the entries in the row of the inverted covariance matrix:

$$(\overset{\wedge}{\sum})^{-1} = \begin{bmatrix} c_{11}....c_{1K} \\ ..... \\ ..... \\ c_{K1}....c_{KK} \end{bmatrix}$$

$$w_k = \sum_{l=1}^{K} c_{kl}$$

(v) Finally create a new variable, $\hat{y}_i$, that is a weighted average of $\tilde{y}_{ik}$ for person i. When constructing $\hat{y}_i$, weight its inputs, standardized variables $\tilde{y}_{ik}$ by the inverse of the covariance matrix of the transformed variables. A simple way to do this is to set the weight on each outcome equal to the sum of its row entries in the inverted covariance matrix for area. The index variable $\hat{y}_i$ is called because this transformation yields a generalized least squares estimator Anderson (2008) .

$$\hat{y}_i = \Big(\sum_{k \in K} w_k \Big)^{-1} \sum_{k \in K_i} w_k * \frac{y_{ik} - \bar{y_k}}{\sigma_k^y}$$

F Learning Assessment

Figure F.4: Learning Assessment

(TEST INFO) SCHOOL NAME: _____ DATE: (DD)___/(MM)___/(YY)___ SURVEYOR ID: ____
(PUPIL INFO): FIRST NAME: _____ LAST NAME: _____ GRADE AND CLASS: _____
COMMENTS:

*For each question below, **four answers are given, but only one of the four is right**. Work out which is the best answer.*
*Then, **TICK** the answer of your choice. For example, if you had chosen answer B for a question, you would show it like this:* ☒ **B.**

1) Which animal is represented in the image?

- ☐ A. Donkey
- ☐ B. Snake
- ☐ C. Monkey
- ☐ D. Butterfly

2) 11 + 13 =
- ☐ A. 14
- ☐ B. 24
- ☐ C. 25
- ☐ D. 36

3) Which day comes before Friday and after Wednesday?
- ☐ A. Monday
- ☐ B. Tuesday
- ☐ C. Thursday
- ☐ D. Saturday

4) How many triangles are on the picture?

- ☐ A. 39
- ☐ B. 42
- ☐ C. 49
- ☐ D. 54

5) Did you see the man ..... stole the car?
- ☐ A. who
- ☐ B. which
- ☐ C. whom
- ☐ D. when

6) 18 ÷ 2 =
- ☐ A. 6
- ☐ B. 8
- ☐ C. 9
- ☐ D. 12

7) When he was young, Lawrence used to ..... a lot with his sisters?
- ☐ A. played
- ☐ B. plays
- ☐ C. play
- ☐ D. playing

8) 11 x 10 =
- ☐ A. 100
- ☐ B. 101
- ☐ C. 110
- ☐ D. 110

9) Choose the picture best describing the image:

- ☐ A. The lady is wearing a hat
- ☐ B. The lady has long hair
- ☐ C. The lady is raising her arm
- ☐ D. The lady is eating

10) 21 ☐ 3 = 7
- ☐ A. x
- ☐ B. ÷
- ☐ C. -
- ☐ D. >

11) Mercy is ..... than her sister
- ☐ A. Younger
- ☐ B. More young
- ☐ C. Youngest
- ☐ D. Most young

12) What is the denominator of the fraction $\frac{3}{6}$
- ☐ A. 9
- ☐ B. 3
- ☐ C. 6
- ☐ D. 18

13) Gift is a hairdresser. Which picture represents Gift?

- ☐ A. Picture 1
- ☐ B. Picture 2
- ☐ C. Picture 3
- ☐ D. Picture 4

14) Moses had 6 mangos and ate 2 of them. How many mangos remain?
- ☐ A. 2 mangos
- ☐ B. 4 mangos
- ☐ C. 5 mangos
- ☐ D. 8 mangos

15) Martin was born in 2003.
- ☐ A. He is 14 years old.
- ☐ B. He does 14 years.
- ☐ C. He has 14 years.
- ☐ D. He measures 14 years old.

16) She laughs because the joke is:
- ☐ A. Sad
- ☐ B. Funny
- ☐ C. Easy
- ☐ D. Wrong

17) 2,23 + 1,07 =
- ☐ A. 2,30
- ☐ B. 3,27
- ☐ C. 3,30
- ☐ D. 4,13

18) Which one is true?
- ☐ A. Goats eat lions
- ☐ B. Chickens eat humans
- ☐ C. Cats eat mice
- ☐ D. Birds eat crocodiles

19) How many K50 notes are in K500?
- ☐ A. 5
- ☐ B. 10
- ☐ C. 50
- ☐ D. 100

20) Anna is walking home with ..... brother.
- ☐ A. his
- ☐ B. her
- ☐ C. him
- ☐ D. theirs

21) 2410 – 1521 =
- ☐ A. 889
- ☐ B. 756
- ☐ C. 1029
- ☐ D. 999

22) Choose the sentence describing the image:

- ☐ A. The girl sweeps the floor
- ☐ B. The chicken lays an egg
- ☐ C. The girl collects the eggs
- ☐ D. The eggs hatch

23) Find the perimeter of the shape below:


- ☐ A. 20 cm
- ☐ B. 29 cm
- ☐ C. 38 cm
- ☐ D. 40 cm

24) Cheelo woke up late ..... he missed the bus.
- ☐ A. but
- ☐ B. so
- ☐ C. for
- ☐ D. because

25) Identify two lines which are parallel:


- ☐ A. Lines AB and BC
- ☐ B. Lines AC and DB
- ☐ C. Lines AD and BC
- ☐ D. Lines AD and CD

26) Which of the following is correct?
- ☐ A. There are five days in a week
- ☐ B. There are sixty seconds in a minute
- ☐ C. There are twenty days in a month
- ☐ D. There are fifteen months in a year

G Attention Test

Figure G.5: Learning Assessment